The complement coercion phenomenon: Implications for models of sentence processing

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by

Yao-Ying Lai

Dissertation Director: Maria Mercedes Piñango

May 2017

Contents

CONTENTS

List of Figures

List of Tables

Acknowledgments

Many things happened in my years as a graduate student, and they were quite dramatic as well as life-changing to me. During the period, I have received help and support from a lot of people, whose generosity and kindness I will never forget.

Firstly, I want to thank my advisor, Maria Piñango, who has guided me through every study that I have done and helped me out from all the academic or personal struggles in my time at Yale. I am always impressed by her insightful thoughts and those inspiring discussions in her seminars and our meetings. In addition, Maria was extremely supportive when I was ill; without her considerations and warm encouragement I would not survive this Ph.D study. Although I have said "thank you" for many times, my gratitude to her is really beyond words.

I am fortunate to have an excellent thesis committee: Ashwini Deo, David Braze, and Jason Shaw.

I would like to thank Ashwini Deo for her advice on the studies and for asking critical questions on my research work, which encouraged me to sharpen

my thoughts and clarify important issues. Also, Ashwini's seminars are among the most insightful courses that I have taken.

David Braze has been very patient with me and he always made time to teach me the eye-tracking techniques, answer my questions, and give me constructive suggestions. I have learned a lot from him in the process, including skills in R, logical thinking, efficient troubleshooting, and staying calm (as he always is).

I am grateful to Jason Shaw, who kindly agreed to join my thesis committee. In these past months, he gave me useful and encouraging feedback on the chapters, and shared his thoughts on the studies.

The faculty members of the department have helped me from time to time throughout these years. I have worked with Bob Frank and Jim Wood as a teaching fellow, and would like to thank them for being great instructors of the courses. A special thank goes to Steven Anderson, who advised my qualifying paper on morphology with encouragement and kind words. I also want to thank Jelena Krivocapić, who was willing to point out my weakness straightforwardly for me to improve in my first semester as a graduate student.

I would like to thank the members of the Language & Brain Lab for their help on my research, and Cheryl Lacadie for teaching me the fMRI analyses with patience. During the past year, I spent quite some time working in Haskins Laboratories. I was lucky to be surrounded by many knowledgeable and extremely nice people there, including (but not limited to) Susan Shuai,

Gong Tao, Jenna, Yi-Hui, and Tammy, who have helped me with experiment implementation and data analyses.

My friends are indispensable in my years studying abroad. My sincere thanks go to Tzuyu Kao & John Chu, Li-Fang Chang & Shih-Chieh Ti, who have become like my family in the U.S., supporting me whenever I need. Also, I want to thank Jhen-Si Chen, Rebecca Liu, Angela Hsieh, Vicky Kao, and Sophie Lin, for their accompany during my time in New Haven.

My special gratitude goes to Ting-Wen Lan, who has accompanied me walking through every up and down in the past ten years. I appreciate all the joy and happiness that he has brought to me.

I owe my family too much, especially Jui Hung, Yi-Hsuan Hung, Zheng-Xin Lai, Chi-Han Chang, and Yu-Hsuan Chang, who constantly send their love from the other side of the earth and comfort me whenever I am depressed.

Finally, the person I am most grateful to, but have difficulty finding words to convey the depth of my gratitude, is my dear mother—Jung-Wei Hung. She has always been my superwoman and harbor. Her infinite love, courage, kindness, perseverance, tolerance and wisdom are the pillars of my life.

— To my mother, my family, and Ting-Wen

Chapter 1

Introduction

1.1 Overview

In this dissertation, I ask a twofold research question: (1) How to understand the complement coercion phenomenon, and (2) how do comprehenders obtain an underspecified sentence meaning in real-time processing? These two questions are connected because sentences argued to exhibit a pattern of complement coercion instantiate a case of meaning underspecification associated with additional processing cost. In sentences involving complement coercion, there is a piece of meaning that is not morpho-syntactically supported yet obtained by comprehenders (see Section 1.2).

To answer the research questions, I examine different accounts that are proposed to capture the complement coercion phenomenon and the associated

processing cost via four studies. I aim to integrate the understanding of this phenomenon into a processing model based on the empirical findings. More specifically, I examine how lexical representation, semantic composition, and context constraints factor in real-time comprehension. These are necessary components for comprehending sentences involving semantic underspecification and should be taken into consideration by any processing model.

This work connects two bodies of literature: (a) linguistic analyses of complement coercion and the hypotheses for its processing, and (b) views of real-time sentence comprehension, including a "mismatch-repair" approach and a constraint-satisfaction approach. While most previous models focus on the syntax-semantics interface, investigating how and when semantic factors influence sentence processing, this dissertation targets the processing of semantic representation while syntactic structure remains constant, bypassing the complication of syntactic factors. By focusing on semantic processing that is dissociable from syntactic structure, this dissertation contributes to uncover how comprehenders obtain meaning that is not morpho-syntactically supported in real-time.

In Section 1.2, I briefly sketch the complement coercion phenomenon and describe two hypotheses that capture the processing of complement coercion. In section 1.3, I provide a preview of my claims. Section 1.4 outlines the content of each chapter.

1.2 The complement coercion phenomenon and its processing

Complement coercion refers to the phenomenon that an entity-denoting complement receives an eventive interpretation when composed with certain verbs such as *begin* and *enjoy* (Jackendoff, 1997; Pustejovsky, 1991, 1995). For example, sentences such as (1.1a) often receive an interpretation as in (1.1b).

(1.1) a. *The woman began/enjoyed the book.*

b. $\rightsquigarrow$ *The woman began/enjoyed* <u>*reading/writing/etc.*</u> *the book.*

The entity-denoting complement (*the book*) is coerced into an event in (1.1b), and the verbs that trigger complement coercion are called "coercion verbs." Noticeably, the underlined event associated with the complement in (1.1b) is not overtly stated in (1.1a) and cannot be derived by any individual lexical item in the sentence. This demonstrates a phenomenon of semantic composition that is isolable from syntactic structure.

In terms of real-time comprehension, a series of studies have shown that processing sentences with complement coercion like (1.1a) engenders additional cost and brain activity (Baggio et al., 2010; Delogu and Crocker, 2012; Frisson and McElree, 2008; Husband et al., 2011; Kuperberg et al., 2010a; Lapata et al., 2003; McElree et al., 2001, 2006; Pickering et al., 2005; Piñango and

Zurif, 2001; Pylkkänen and McElree, 2007; Pylkkänen et al., 2009; Traxler et al., 2002, 2005). For instance, McElree et al. (2001) found that the (1.2a) engenders longer reading times than (1.2b).

(1.2) a. *The student began the book late in the semester.*

b. *The student read the book late in the semester.*

— Example from McElree et al. (2001)

These observations lead to the following questions: (i) How do comprehenders obtain the unstated meaning in sentence (1.1a)? (ii) What is the source of the additional processing cost and brain activity associated with sentences involving complement coercion? There are a few analyses that attempt to capture the unstated sentence meaning, and different analyses lead to different processing hypotheses that account for the cost associated with complement coercion. These are discussed in detail in Chapter 2; below I briefly sketch two hypotheses as a preview.

On the traditional view, coercion verbs, such as *enjoy* (psychological verbs) and *begin* (aspectual verbs), are assumed to select for event-denoting complements. The composition of an event-selecting coercion verb and an entity-denoting complement gives rise to a mismatch of semantic type due to the violation of the verb's selectional restriction. The mismatch calls for a type-shifting operation that coerces the complement from an entity to an event, in order to satisfy the selectional restriction of the verb (Jackendoff, 1997;

Pustejovsky, 1991, 1993). On this view, the processing cost associated with complement coercion results from the insertion of the type-shifting operation in comprehension to repair the type mismatch between the verb and its complement (e.g., McElree et al., 2001; Pickering et al., 2005; Pylkkänen and McElree, 2007; Traxler et al., 2002). I will call this treatment for the extra processing cost *the Type-Shifting Hypothesis*.

An alternative account for the complement coercion phenomenon questions the assumption of the Type-Shifting Hypothesis, pointing out that the complement of a subset of coercion verbs, namely aspectual verbs, does not always receive an eventive reading. Unlike the Type-Shifting Hypothesis which targets coercion verbs as a whole, Piñango and Deo (2012, 2015) propose the Structured Individual analysis that targets aspectual verbs (e.g. *begin, finish*). They argue that aspectual verbs select for *structured individuals*, i.e., entities that can be mapped onto an axis, as their complements (see Chapter 3). Sentences with aspectual verbs specify a subpart relation between the participants involved in the sentence. These semantic properties pertaining to aspectual verbs dissociate them from others such as psychological verbs (e.g., *enjoy*, *prefer*).

Piñango & Deo indicate that a sentence with an aspectual verb gives rise to a semantic ambiguity among multiple readings. For example, sentence (1.3) yields an agentive reading, in which the subject referent is perceived as an agent performing some activity associated with the complement denotation, as

in (1.4a). The sentence also yields a constitutive reading, in which the subject referent is perceived as a subpart of the entity denoted by the complement, as in (1.4b), in addition.

(1.3) *Jane Austen began the novel.*

(1.4) a. ⇝ Jane Austen began reading/writing/etc. the novel.

b. ⇝ Jane Austen's excerpt was the initial subpart of the novel.

Piñango & Deo's analysis leads to the Structured Individual Hypothesis (SIH) for the processing of aspectual verbs. The SIH states the observed cost that is considered as reflecting complement coercion actually reflects two processes associated with the processing of aspectual verbs: (a) exhaustive retrieval of the verb's lexical representation, and (b) ambiguity resolution among multiple semantic representations in context. Details of the Structured Individual analysis and hypothesis are presented in Chapter 3.

Notice that although both the Type-Shifting and Structured Individual hypotheses attempt to account for meaning that does not have a morphosyntactic reflex, the type-shifting operation proposed by the Type-Shifting Hypothesis is invoked to overcome mismatches of lexical semantic requirements, not to deal with meaning underspecification *per se* as a formal mechanism. In contrast, the Structured Individual Hypothesis is a proposal to capture how comprehenders obtain a specific semantic representation among multiple alternatives, without involving a repairing operation. The SIH argues that the comprehen-

sion of sentences with aspectual verbs like (1.3) involves ambiguity resolution for semantic underspecification, rather than a mismatch of lexical selectional restriction as suggested by Type-Shifting Hypothesis.

It is important to adjudicate between the Type-Shifting and Structured Individual hypotheses because they point to potentially mutually exclusive architectures of the processing system. The Type-Shifting Hypothesis suggests a *mismatch* → *repair* mechanism: When a representation pursued by the processor contains a type mismatch between the verb and the complement, the processor invokes a repairing operation to reconfigure the semantic representation—a process that incurs computational cost. This hypothesis implies a *semantic garden path* view of meaning composition that involves a misalignment of lexical semantic requirements followed by a repairing mechanism.[1]

In addition, the Type-Shifting Hypothesis is associated with an operational approach that aims to identify a set of semantic compositional operations which interface with the system of language processing (Pylkkänen and McElree, 2006; Pylkkänen, 2008). On this view, the interpretive operations are recruited by the processor during comprehension, and researchers should

[1]The term "semantic garden path" is used in light of the garden path models for local syntactic ambiguities (e.g., Clifton et al., 2003; Ferreira and Clifton, 1986; Frazier and Fodor, 1978; Frazier, 1987). In these models, the processor forms an initial parse based on syntactic strategies; if this initial analysis gives rise to an error, the processor reanalyzes the sentence by recruiting other resources (e.g. world knowledge) to repair. The Type-Shifting Hypothesis does not address syntactic ambiguity; yet it implicates a mismatch-repair mechanism to deal with the processing of semantic representation. In this sense, the term "semantic garden path" is used here to characterize the processing system implied by the Type-Shifting Hypothesis. I am grateful to my thesis committee for suggesting this term.

attempt to figure out what those operations are, and why (or why not) some of them are costly (Pylkkänen and McElree, 2006). Proponents of this operational view suggest that "there is a strong connection between the representational complexity of linguistic expressions and processing difficulty" (Brennan and Pylkkänen, 2010). As Phillips (2013) mentions, this view on complement coercion echoes the Derivational Theory of Complexity (DTC) in attributing the processing cost to an individual operation in the processing system (Fodor et al., 1974, p.210). Postulating an additional operation recruited by the processing system seems to imply a dissociation between the grammar (static linguistic knowledge that is task-neutral) versus the processor (a set of procedures used in real-time by exploiting the grammar)—a separation that may not be necessary (Jackendoff, 2002, 2007, 2010; Lewis and Phillips, 2015; Phillips, 2013).

In contrast, on the Structured Individual Hypothesis, the processor builds multiple semantic representations in parallel, without committing to a single representation. The alternative representations are evaluated by contextual constraints, and unsupported alternatives are pruned. The additional processing cost results from determining an appropriate interpretation among multiple alternatives based on contextual information. The processing system implied by the SIH is strongly lexically rooted and contextually constrained. That is, sentence meaning is constructed by the composition of lexical representations. The composition may give rise to multiple semantic representations,

resulting in a semantic ambiguity. The processor seeks to find a representation that satisfies both lexical and contextual constraints. These processing properties are connectable to the constraint-satisfaction approach such as Crain and Steedman's (1985) model. It (i) allows the processor to construct multiple representations simultaneously, (ii) makes use of contextual constraints to evaluate the alternative representations, and (iii) attributes additional processing cost to the competition among the alternatives. The implications of the Structured Individual Hypothesis are discussed in Chapter 9.

The Type-Shifting Hypothesis and the Structured Individual Hypothesis as well as their implicated processing systems are summarized in Table 1.1.

Table 1.1: Type-Shifting Hypothesis & Structured Individual Hypothesis and their implications for the processing system

Hypothesis	**Type-Shifting**	**Structured Individual**
Processing mechanisms	Mismatch-repair	Constraint-satisfaction
Source of processing cost	Interpolation of a repairing operation	Determination among alternative representations by contextual constraints.

Since different hypotheses for capturing the complement coercion phenomenon implicate different views regarding the processing mechanisms, investigating the source of the cost associated with this phenomenon informs how human processor works to obtain sentence meaning in real-time.

1.3 Forecasting the processing system implicated by the findings of the dissertation

I report four studies (Chapter 4 ∼7) that adjudicate between the Type-Shifting Hypothesis and the Structured Individual Hypothesis. To forecast, results are taken to support the Structured Individual Hypothesis (Chapter 8). I argue that the complement coercion phenomenon should be viewed as the processing of aspectual verbs (which are dissociated from psychological verbs), and that the additional processing cost results from (a) exhaustive lexical retrieval and (b) ambiguity resolution by contextual constraints. Because the Structured Individual Hypothesis takes into account meaning underspecification, it helps to reveal how comprehenders obtain unstated meaning in real-time.

The Structured Individual Hypothesis for meaning underspecification is consistent with the constraint-satisfaction approach for local syntactic ambiguity (Chapter 9). I suggest that a sentence is processed incrementally as described below. When encountering the lexical items of the sentence, the processor retrieves lexical representations exhaustively. In addition, lexical semantic requirements (e.g., selectional restrictions) must be fulfilled within

the sentence. Based on the lexical representations, the processor composes multiple semantic representations in parallel. The alternative representations are narrowed down by contextual constraints, such that the representations unsupported by context are pruned. In the case of semantic underspecification, processing cost results from (a) the composition of multiple semantic representations and (b) the determination of a contextually appropriate representation among multiple alternatives. Additional repairing operation for comprehension is not necessary, as no mismatch is involved in sentences with underspecified meaning.

On this view, sentence processing is rooted in lexical representations (without resorting to operations external to lexical representations, such as type-shifting) and constrained by contextual information, as suggested by the constraint-satisfaction approach (e.g., Crain and Steedman, 1985; MacDonald et al., 1994a,b; McRae et al., 1998; Pearlmutter et al., 1994; Seidenberg et al., 1982; Tanenhaus et al., 2000; Trueswell, 1996). Meanwhile, the Structured Individual Hypothesis is consistent with an architecture in which semantic representation can be established by lexico-semantics without necessarily relying on syntactic structure, as in the Parallel Architecture proposed by Jackendoff (1997, 2002, 2007).

1.4 The structure of the dissertation

Chapter 1 describes the research questions, goals, and contributions of this dissertation, along with the approach taken to answer the research questions. A brief overview of the dissertation is provided.

Chapter 2 characterizes the complement coercion phenomenon, reviews various linguistic analyses that attempt to capture this phenomenon, and presents several hypotheses that account for the associated processing cost.

Chapter 3 elucidates the Structured Individual analysis proposed by Piñango and Deo (2012, 2015) that captures the lexical semantics of aspectual verbs, which are dissociated from other coercion verbs. This analysis leads to the Structure Individual Hypothesis, which claims that processing sentences with aspectual verbs involves exhaustive lexical retrieval and ambiguity resolution in real-time comprehension.

Chapter 4 presents Study 1 (self-paced reading, fMRI) that examines the behavioral and neurological patterns of sentences with aspectual verbs. Results show that aspectual verbs exhibit a processing profile distinct from psychological verbs, while the two verb types are

collapsed under the set of coercion verbs in previous studies on complement coercion.

Chapter 5 presents Study 2 (self-paced reading, fMRI), which shows that all sentences with aspectual verbs, regardless of subject animacy, engender a similar processing pattern in behavioral and neurological measurements.

Chapter 6 reports Study 3 (eye-tracking), which examines the role of context in real-time comprehension of aspectual-verb sentences. Results show that context type (biasing vs. neutral) does not affect the processing of aspectual-verb sentences. This suggests that the processor composes multiple semantic representations from the verb and its complement regardless of context.

Chapter 7 reports Study 4 (questionnaire), which investigates whether a disambiguating context influences the ultimate interpretation and acceptability of sentences with aspectual verbs. Results show that comprehenders arrive at the interpretation supported by the context and that a disambiguating context increases sentence acceptability.

Chapter 8 adjudicates between the Type-Shifting Hypothesis and the Structured Individual Hypothesis based on the results gathered from

Study 1∼4. I argue that the findings are consistent with the Structured Individual Hypothesis and inconsistent with the Type-Shifting Hypothesis.

Chapter 9 discusses the implications of the findings for the processing system. I argue that the processor builds multiple semantic representations in parallel without committing to only one representation. The alternative representations are evaluated by contextual constraints, and non-supported representations are pruned in the course of comprehension. The comprehension of meaning underspecification can be captured by the constraint-satisfaction approach of sentence processing.

Chapter 10 summarizes this dissertation and provides future directions.

Chapter 2

The Complement Coercion Phenomenon and Hypotheses on its Processing

2.1 Coercion: Enriched composition and syntax-semantics mismatch

The Fregean principle of compositionality states that the interpretation of an expression is obtained by meanings of lexical items in the expressions and the way they are syntactically combined (Frege, 1892). However, compositionality is not always obeyed. Some sentences contain a piece of meaning that is not expressed by morpho-syntactic means — neither by the word's lex-

ical semantics nor the syntactic structure — but it naturally arises from the semantic composition.

For example, the sentence "*The student began the paper*" is often read as "*The student began reading/writing the paper.*" The activity associated with the complement is not provided by any individual lexical item of the sentence, nor does the syntactic structure supply any clue for such meaning. Yet the eventive interpretation is obtained during comprehension. The exact activity (e.g. reading, writing, editing) is underspecified and left for comprehenders to determine in context.

In such cases, the meaning is transparent in the semantic representation but blind to the syntactic structure. This is what Jackendoff (1997, p.49) termed **enriched composition**, in contrast to *transparent* or *simple composition*, which obeys strict compositionality in that all meanings are expressed morphosyntactically. The semantic composition is "enriched" in the sense that it yields an "extra" meaning that is not rendered by function application following syntactic derivation. Enriched composition therefore violates strict compositionality, which assumes a one-to-one mapping between the semantic and syntactic representations. The underspecified meaning in enriched composition is left to comprehenders to reconstruct in order to make sense out of the sentence (Culicover and Jackendoff, 2005, p.228).

The example instantiating enriched composition above involves **coercion**, in which an underspecified meaning of the sentence is obtained by changing

some semantic feature of a lexical item in order to achieve well-formedness in the semantic representation (Pustejovsky, 1989, 1991; Jackendoff, 1997; Culicover and Jackendoff, 2005). Coercion encompasses a range of phenomena, including Aspectual Coercion, Mass-count Coercion, and Complement Coercion. I will briefly introduce the first two below, and delve into complement coercion, which is the focus of this dissertation, in Section 2.2.

Aspectual coercion. Sentences like (2.1a) usually give rise to an iterative interpretation, as in (2.1b).

(2.1) a. John jumped for an hour.

b. John jumped [REPEATEDLY] for an hour.

However, the iterative meaning is not expressed morphosyntactically by lexical items or the syntactic structure of (2.1a). Rather, it is carried out in the semantic representation, in which the punctual meaning of the predicate (a single jump) is interpreted as an iterative one (a series of jumping events). Notice that semelfactive predicates (e.g. *jump, blink*) by definition describe punctual, bounded situations, and are incompatible with modifiers which select for atelic, durative predicates. The composition of a semelfactive predicate and a durative *for*-adverb hence results in an aspectual mismatch. However, instead of judging such sentences as anomalous, comprehenders make sense out of them by enriching the semantic representation. It is claimed that the aspectual mismatch between the punctual predicate and the durative adverbial phrase is

resolved by an operation that turns the punctual reading of the semelfactive predicate into an repetitive one, to ensure semantic well-formedness (Jackendoff, 1997, p.52). This phenomenon is called *Aspectual Coercion* (Jackendoff, 1997; Moens and Steedman, 1987; Pustejovsky, 1991, 1995).

Mass-Count coercion. Using a count noun (e.g., *chicken*) as a mass noun shifts its reference from denoting a count singular bounded object to its substance (e.g., edible parts of the object). For instance, (2.2a) yields the reading in (2.2b).

(2.2) a. He had chicken for dinner yesterday.

b. He had [chicken MEAT] for dinner yesterday.

Reversely, pluralizing a mass noun shifts its meaning from denoting substance to count singular bounded object, as in (2.3).

(2.3) I'd like to have another coffee.

It is suggested that the meaning change is accomplished via a semantic operation that coerces the [Mass: substance] $\longleftrightarrow$ [Count: bounded object] denotations bidirectionally. In (2.3), it is the "coerced" meaning that licenses the use of the determiner *another* (Jackendoff, 1991, Jackendoff, 1997:57; Culicover and Jackendoff, 2005:229)

As an instantiation of enriched composition[1], coercion evidences a **syntax-**

[1]In addition to coercion, another case of enriched composition is reference transfer (Jackendoff, 1997). It refers to the phenomenon that, using the example from Nunberg (1979),

semantics mismatch in that the mapping between the semantic composition and the syntactic derivation is not one-to-one. This suggests that semantics and syntax are independent to each other, contrary to the syntactocentric view that semantic representation derives from syntactic structure necessarily. Investigating coercion therefore allows for addressing an aspect of semantic representation that is isolatable from syntax, providing a chance to examine how semantic interpretation is obtained in the absence of syntactic constraints.

In what follows, I turn to the target of this dissertation—complement coercion. I will present the phenomenon (Section 2.2), linguistic analyses (Section 2.3), and experimental work investigating its processing profiles (Section 2.4).

2.2 Complement coercion and the eventive interpretation

It has been observed that sentences combining verbs such as *begin* and *enjoy* with an entity-denoting complement, as in (2.4a), often give rise to an eventive reading, as in (2.4b). The phenomenon that an entity-denoting complement receives an eventive interpretation is called *complement coercion* (Jackendoff, 1997; Pustejovsky, 1991, 1995).

(2.4) a. John *began/enjoyed* the book.

the subject in "*The ham sandwich in the comer wants some more coffee.*" is interpreted as "a person contextually associated with the ham sandwich."

b. John *began/enjoyed* reading/writing/etc. the book.

Notice that the eventive interpretation is not expressed lexically or syntactically but is obtained through the composition of the lexical items in the semantic representation — a case of *enriched composition* as mentioned in Section 2.1 (Jackendoff, 1997, p.49).

For cases like (2.4), a question naturally arises is how readers arrive at the interpretation (2.4b) with an unstated eventive interpretation (regardless of the specific activity, such as reading, writing, or burning). Below I discuss several linguistic analyses on how this underspecified meaning is obtained in sentences involving complement coercion.

2.3 Linguistic analyses of complement coercion

In literature, complement coercion has received a few treatments (*cf.* Pylkkänen and McElree, 2006). A long-standing approach is the Type-Shifting analysis, which asserts that complement coercion involves a type-shifting operation that coerces the entity-denoting complement to an event (Jackendoff, 1997, 2002; Pustejovsky, 1991, 1995). The unexpressed event associated with the complement is obtained via

— selecting the contextually appropriate quale from the qualia structure

of the complement noun (Pustejovsky, 1991, 1995)

— interpolating an interpretive function at the Syntax Structure (SS)-Conceptual Structure (CS) interface and looking for the target event (Jackendoff, 1997, 2002)

Another approach claims that complement coercion involves a hidden VP structure with an empty Verb head, which is saturated by pragmatical inference in context (de Almeida, 2004; de Almeida & Dwivedi, 2008). These analyses are discussed in detail in Section 2.3.1 and 2.3.2 respectively.

2.3.1 The Type-Shifting analysis

The traditional approach for complement coercion assumes that verbs such as *begin* and *enjoy* impose a selectional restriction on their complement, requiring it to denote an event (Briscoe et al., 1990; Pustejovsky, 1991, 1995; Pustejovsky and Bouillon, 1995). Based on this assumption, combining an event-selecting verb with an entity-denoting noun, as in (2.4a), results in a **type mismatch** in the semantic representation. This mismatch calls for a **type-shifting** operation that coerces the semantic type of the complement from an entity to an event, in order to satisfy the selectional restriction of the verb. The verbs that trigger complement coercion are referred to as *coercion verbs* accordingly. This approach will be called *the Type-Shifting analysis*.

2.3.1.1 Type-Shifting as a repairing operation for semantic mismatch

The notion of type-shifting originates as a means to resolve a type mismatch between two phrases in conjunction (Partee and Rooth, 1983; Partee, 1987). In principle, the coordinated phrases in conjunction should be of the same semantic type. On this assumption, sentence (2.5) is problematic because a proper name denoting an individual "*John*" of type ⟨e⟩ is conjoined with a quantified NP "*every guest*" of type ⟨⟨e, t⟩ ,t⟩.

(2.5) [John] and [every guest] had beer at the party.

This type mismatch can be resolved by type-shifting *John* from type ⟨e⟩ to the type of the quantified NP, ⟨⟨e, t⟩ ,t⟩, to make the conjuncts consistent in semantic type.

Another motivation for type-shifting is the case in which an NP is conjoined with an adjective phrase (AP). Consider (2.6) from Partee (1987).

(2.6) Mary considers John [competent in semantics] and [an authority on unicorns].

Because the conjoined constituents must have identical type by assumption and because APs are of the type ⟨e, t⟩, the NP "*an authority on unicorns*" must be of the type ⟨e, t⟩ as well. This is achieved by a type-shifting operation that converts the indefinite NP to the type consistent with its conjoined phrase

in the predicative position, i.e. $\langle\langle e, t\rangle, t\rangle \rightarrow \langle e, t\rangle$. With the two phrases now being of the same type, they can be conjoined without a type mismatch.

The mechanism of type-shifting is introduced as a modification to Montague Grammar (Montague, 1973), which requires a correspondence between syntactic categories and semantic types. To ensure direct compositionalty—that each step of semantic composition (such as function application) maps to a step of syntactic derivation— Montague suggests to assign the *highest* semantic type needed for to a given syntactic category, in order to avoid any type clash. For sentence (2.5) to be semantically well-formed by using compositional rules that correspond to syntactic rules, Montague's strategy lists all NPs as of the type $\langle\langle e, t\rangle, t\rangle$ to allow conjoining an NP of an individual and a quantified NP. Partee and Rooth (1983), while preserving the category-type mapping, allow a given syntactic category to have a family of possible semantic types. In contrast to Montague, Partee and Rooth assign each item the *simplest* type ("all expressions are interpreted at the lowest type possible"), and invoke a type-shifting operation to convert the simple type to a higher one when type coherence is required.

This treatment is adopted by Pustejovsky (1989, 1993). Consider his example (2.7a), which is interpreted as (2.7b).

(2.7) a. John considers Mary [a fool].

b. John considers Mary to be a fool.

Pustejovsky suggests that verbs select for an argument of a particular type. Assuming that the verb *consider* selects for a predicative phrase, following Partee (1987), he argues that the NP "*a fool*" in (2.7a) is shifted to the type of a predicate, ⟨e, t⟩.

For Pustejovsky (1993, 1995), type-shifting is "licensed" by a governing verb that imposes selectional restriction on its argument. The type of a syntactic category is assigned a uniformed type but may be realized differently as requested by a governing function in context. For example, the complement of the verb "*want*" appears in several different forms (Pustejovsky, 1995, p.110).

(2.8) a. Mary wants John to leave. (S [+INF])

b. Mary wants to leave. (VP [+INF])

c. Mary wants a beer. (NP)

Suppose that the verb "*want*" selects for a proposition as the type of its argument. If the complement of the verb matches such type, as in (2.8a), the composition is well-formed. However, if the complement does not denote a proposition, as in (2.8b) and (2.8c), then the complement is coerced to the appropriate type as required by the verb. According to Pustejovsky, semantic composition makes reference to the type-shifting operation to fix any type mismatch (1995: 111). The rule of functional application with coercion is given in (2.9) below.

(2.9) **Function Application with Coercion (FAC)**: if α is of type c, and

β is of type $\langle$a, b$\rangle$, then,

(i) if type c = a, then $\beta(\alpha)$ is of type b.

(ii) if there is $\sigma \in \Sigma_\alpha$ such that $\sigma(\alpha)$ results in an expression of type a, then $\beta(\sigma(\alpha))$ is of type b.

(iii) otherwise a type error is produced.

—Pustejovsky (1995, p.110)

To recapitulate, in Pustejovsky's framework, a function denoted by a verb selects for a particular semantic type as its argument. When this requirement is not met, coercion is called for to type-shift the argument to the type expected by the verb. This, as Pustejovsky suggests, accounts for the phenomenon of complement coercion.

2.3.1.2 Pustejovsky (1991, 1995): Type coercion and Qualia structure

Pustejovsky (1995, p.115) presents the following paradigm:

(2.10) a. John began [a book].

b. John began [reading a book].

c. John began [to read a book].

(2.11) a. Mary enjoyed [the movie].

b. Mary enjoyed [watching the movie].

He suggests that the complement of verbs such as *begin* and *enjoy* denotes an event as its semantic type, regardless of its syntactic form. In (2.10a) and (2.11a), the complements (*a book*, *the movie*) denote entities, incompatible with the selectional restriction of their governing verbs. The composition of the entity-denoting complement and the event-selecting verb results in type mismatch in semantic representation. This calls for **Type Coercion**, defined by Pustejovsky (1991) as "a semantic operation that converts an argument to the type that is expected by a function, where it would otherwise result in a type error." Coercion requests that the complement must belong to the correct type expected by the verb so as to avoid a semantic type mismatch during composition. By FAC in (2.9), the complements must be type-shifted to an event to ensure semantic well-formedness.

While an event type is in place for the complement after type coercion, one would ask how the exact activity associated with the complement is obtained. In Pustejovsky's (1991, 1995) *Generative Lexicon*, the event associated with the complement denotation is reconstructed from the **Qualia Structure**—a system that characterizes the semantics of a nominal in the following four aspects:

(2.12) a. *Telic Role* specifies its purpose and function.

b. *Agentive Role* specifies whatever brings it about.

c. *Constitutive Role* specifies the relation between it and its con-

stituent parts.

d. *Formal Role* specifies the basic category of which distinguishes it within a larger domain.

Each qualia role can be viewed as "partial functions (Q_T, Q_A, Q_C, Q_F) from a noun denotation into its subconstituent denotations," returning the value of a particular qualia role (Pustejovsky, 1991). With respect to complement coercion, the relevant ones are the Telic and Agentive Role. In "*John began the book,*" the function Q_T(book) returns the interpretation of *reading the book* (its purpose and function), and Q_A(book) returns the interpretation of *writing the book* (how it is brought about).

That is, in "*John began the book*," the eventive reading (e.g. *John begin reading/writing the book*) is obtained by applying the coercion operation on the complement, type-shifting it to an event type, and the specific event is reconstructed via the Telic or Agentive role in the qualia structure of the complement noun.

Pustejovsky and Jezek (2008) suggest that coercion is accomplished by two processes: First, the verb introduces an event (Event Introduction); then the value of the Qualia encoded in the complement noun is "lifted at the level of interpretation (Qualia Exploitation, where a subcomponent of the argument's type is accessed and exploited)" as requested by the verb. They suggest that the meaning of a word is interpreted compositionally, modulated by context

and semantic operations such as type-shifting.

For Pustejovsky (1991, 1995), coercion manifests *logical metonymy* in that the event—as a subpart in the qualia structure of the complement—selected by the verb stands for the entity-denoting complement. The qualia structure is a way of capturing logical metonymy by linking subconstituent (i.e., qualia roles) encoded in lexical items. A highly organized lexical representation of nominals plays an active role in generating the meaning of a sentence. In complement coercion, the meaning of the sentence is determined not only by functional application of the verb to its argument but also by functional application of the qualia from the complement noun to its subconstituent event denotation. This constitutes a case of what Pustejovsky termed *cocomposition*— a structure which involves not only the function application of the verb to its complement but also a function application from the complement's denotation to its particular qualia role.

2.3.1.3 Jackendoff (1997): Interpretive function at the SS-CS interface

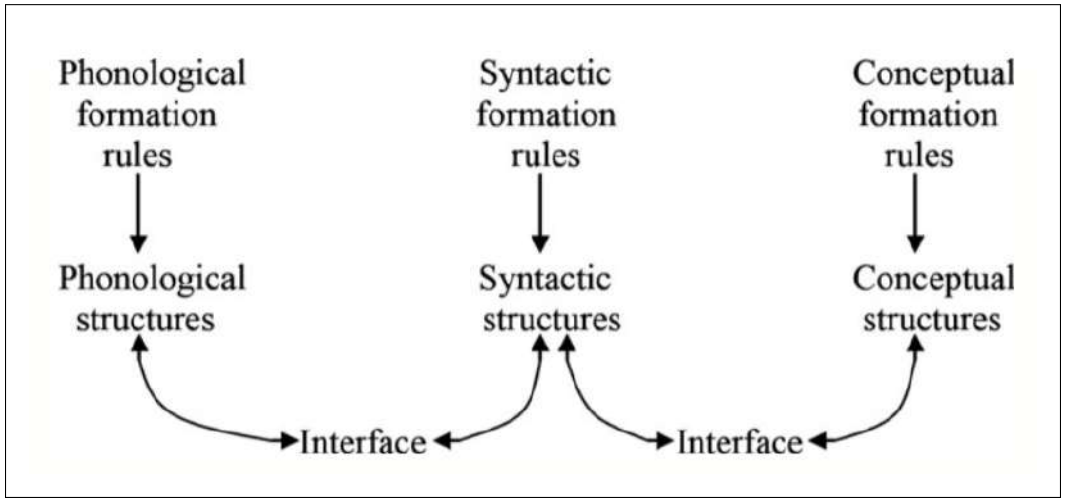
While Pustejovsky captures the unstated eventive meaning by an enriched organization within lexical items, Jackendoff (1997, 2002) suggests that the interpretation is generated by an interpretive operation that lies in the semantic-syntax interface. Jackendoff (1997, 2002, 2010) argues for a tripartite Parallel

Architecture (PA), which consists of three independent modules: Phonological Structure, Syntactic Structure, and Conceptual/Semantic Structure. These structures are connected to each other through a set of interface/correspondent rules, shown in Figure 2.1.

In the PA, semantics is identified as "the organization of those thoughts that language can express," while "thoughts expressed by language are structured in terms of a cognitive organization called *conceptual structure* (CS) (2002: 123)." Conceptual Structure is specified as part of thought, not language. It is where the meanings of linguistic utterances are understood in context, a space where pragmatics or "world knowledge" is integrated in the understanding of sentence meaning. Within this framework, "semantic structure" (as indicated in Pinker (1989) and Bierwisch & Lang (1989)) is the same level of representation as conceptual structure (Jackendoff, 1985: 209), a consequence of the Syntactic Structure (SS)-Conceptual Structure (CS) rule (Jackendoff, 1997: 220, note 8).

In terms of complement coercion, Jackendoff (1997) follows Pustejovsky's account in assuming that the complement of verbs like *begin* and *enjoy* denote an activity (1997: 61). If their complement surfaces as an entity-denoting expression, then two steps ensue. First, the rule in (2.13) interpolates a function that permits the complement NP to denote an activity in satisfying the selectional restriction of the verb. The exact activity is left unspecified at this point.

Figure 2.1: Jackendoff's Parallel Architecture (2010: 3)



(2.13) Interpret NP as [$_{\text{Activity}}$ F(NP)]. (i.e. an unspecified activity involving NP, "doing something with NP")

In the second step, the activity associated with the complement NP is specified. In Pustejovsky's (1991, 1995) lexical decomposition approach (*Generative Lexicon*), this is done by accessing the complement noun's qualia structure —a fine-grained lexical semantic structure (which is different from compositional semantics)[2]. In the Parallel Architecture, the meaning-generating mechanism is located at the interface between the linguistic structure and the non-linguistic cognitive space, namely CS, instead of the qualia structure within the complement noun. For Jackendoff, the specific activity is interpreted in the SS-CS interface, whose correspondence rules permit certain content to be elided from morphosyntax if the meaning is reconstructible from the context (1997: 62).

In comparison, the eventive interpretation is generated by type coercion and exploitation of qualia structure in Pustejovsky's approach but is obtained at the SS-CS interface in Jackendoff's PA, while both assume that coercion verbs semantically select for events as their complements. Different from Pustejovsky, who maintains a semantic-pragmatic dissociation, Jackendoff holds that the two are encompassed in Conceptual Structure in the Parallel

[2]Pustejovsky (1991) claims that there are several distinct semantic levels that contribute the meaning of an utterance in context, such as lexical semantics, compositional semantics, discourse structure, with lexical semantics being just one of them. Each of these levels forms a local interpretation. When integrated, a globally coherent interpretation will be formed "by processes of cooperation among separate semantic models."

Architecture.

2.3.2 Pragmatic inference and the Empty V analysis

The lexical decomposition approach suggested by Pustejovsky (1995) is criticized by Fodor and Lepore (1998). They argue that the Generative Lexicon approach collapses lexical representation, which is taken to be part of linguistic grammar, with "world knowledge", which is not. Also, the interpretation for the complement noun does not necessarily come from the Telic role in its qualia structure. For instance, "*start NP*" does not always mean "start to use NP to perform its function," as "*start a car*" in "*I started a car with a crank.*" does not mean "start to drive a car."

Therefore, the function and use of the object denoted by the complement noun do not need to be specified in its lexical semantics in a decompositional way described by the qualia structure. Instead, the specific activity associated with the complement noun is obtained by *inference* in a later stage of comprehension, dependent upon the comprehender's knowledge about the object's property.

Following Fodor and Lepore (1998), De Almeida (2004) states that the interpretation of sentences involving complement coercion can be obtained via a "late interpretive inferential processes triggered by underspecified con-

structions," instead of the type-shifting operation interpolated in the semantic composition.

De Almeida (2004) reasons that if the type-shifting operation is a lexical process to select the Telic role in the complement noun's qualia structure, then context would either narrow down the range of events (Telic role) specified for the complement or pre-determine the exact event associated with it (Lascarides & Copestake, 1998). In such case, the type-shifting operation does not need to be invoked. In his self-paced reading experiment (de Almeida, 2004, Exp. 2), in which a context sentence precedes the target sentence with an entity-denoting complement NP, de Almeida contrasts three conditions: coercion (*began*), preferred reading (*read*), and non-preferred reading (*wrote*).

(2.14) [Context]: The student was a procrastinator. He was behind schedule and had a lot of work to catch up on for his English class.

[Target]: The student began/ read/ wrote the paper late in the semester.

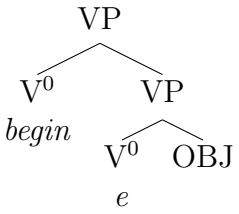
Results show that the coercion and non-preferred-reading conditions patterned alike; both induced longer reading times (RTs) than the preferred reading condition at the post-verb positions. Moreover, in the absence of context, the conditions did not differ.

Based on these results, de Almeida dispenses with the type-shifting op-

eration over lexical-semantic information. He argues that what is reported as the type-shifting effect is a reflex of "a pragmatic process of integrating less preferred constructions within the context." Specifically, expressions like "*begin/enjoy the book*" violate Gricean conversational maxim of Quality (be informative) and Manner (avoid obscurity and ambiguity). According to him, the phenomenon of complement coercion is purely pragmatic in nature, and the interpretation is obtained via "a set of inferential processes beyond the logical form (and truth-condition meaning) of an expression, together with the denotation of its constituent items."

De Almeida and Dwivedi (2008) further flesh out the pragmatic approach, claiming that the [begin NP] construction contains an extra silent VP with an empty verbal head. Syntactically, verbs like *begin* require the structure in (2.15), asking for an event reading.

(2.15)



When followed by an entity-denoting complement NP, as in "*begin the book*," pragmatic inference is called for to saturate the missing activity realized in the empty V head based on contextual information.

Several distributional arguments are provided to support this view. First, VP modifiers scope over the empty V. In their example below, *again* modifies the lower VP *V-ing* (*watching Gandhi*) instead of the higher one (*prefer* VP).

(2.16) We watched *Fargo* last week. I prefer *Gandhi* again.

Second, on the empty V analysis, sentences like "*This book is finishable*" in (2.17a) has the structure in (2.17b), and is interpreted as (2.17c) (Pylkkänen & McElree, 2006).

(2.17) a. [This book] is finishable.

b. $[_{VP}\ [_{V^0}\ \mathbf{e}]\ [_{NP}$ This book]] is finishable.

c. *It is possible to finish* [*V-ing*] *this book.*

Therefore, it is possible for the lower VP to raise to the subject position in the deverbal adjective (V-*able*) structure.

(2.18) $[_{VP}$ Reading this book] is finishable.

Third, while an event-denoting NP complement can alternate between a causative (2.19a) and an unaccusative (2.19b) structure, the event-denoting full VP (gerundive or infinitival) cannot raise to the subject position in an unaccusative structure, as demonstrated in (2.20).

(2.19) a. The students began [the fight].

b. [The fight] began.

(2.20) a. *$[_{VP}$ Reading the book] began.

b. *$[_{VP}$ To read the book] began.

Interestingly, entity-denoting NP complements pattern like event-denoting full VPs; both cannot appear in the subject position of the unaccusative structure.

(2.21) *[The book] began.

This similarity follows from the analysis that the object NP (*the book*) is generated as the complement of an empty V, as [$_{VP}$ [$_{V^0}$ **e**] [$_{NP}$ the book]]. Thus, the event-denoting VP with an empty V behaves like the full VP counterpart; both are disallowed in the subject position.

De Almeida & Dwivedi propose that sentences with event-selecting verbs presuppose a relation z (the verb-referring event)[3] between x (NP1) and y (NP2) such that *x began/enjoyed* ***z-ing*** *y.* In the coercion construction, where a verb is followed by an entity-denoting complement NP (e.g. *begin the book*), this z relation and its associated V head is missing. Comprehenders therefore need to assign a z relation[4] and this is the locus of the processing cost reported as the complement coercion effect. On this view, complement coercion involves no enriched composition nor type-shifting operation; the strict compositionality is therefore preserved.

[3]They state that this z relation is "a semantic variable at the interface between linguistic and conceptual/pragmatic representation."

[4]This proposal implies that an event is always retrieved in the configuration of a coercion verb + an entity-denoting complement,

2.3.2.1 Challenges to the Empty V analysis

Although the empty V analysis for the eventive interpretation in complement coercion accounts for several distributional observations, it is called into question by a few counterarguments pointed out by Pylkkänen and McElree (2006) and Pylkkänen (2008).

The first problem is the scope of adverbial modification. Recall that on the Empty V analysis, coercion sentences involve two VPs as shown in (2.15). Adverbials can modify the lower VP in the overt VP configuration (2.22a), as in (2.22b-c).

(2.22) a. John began writing the paper.

b. John began writing the paper *slowly.*

c. John began writing the paper *on his laptop.*

However, adverbials cannot modify the lower VP of the coercion configuration, as in (2.23). If coercion expressions contain an underlying VP with an empty V, they should allow modifications scoping over the lower VP, contrary to the fact.

(2.23) a. We finished the meal slowly.

→ *False a slow meal comes to an end quickly.*

b. #I started a loaf of bread with a knife.

→ *False if an event of cutting the loaf of bread with a knife was*

initiated. The knife must be an instrument of the initiation.

—Examples from Pylkkänen and McElree (2006)

The second problem is passivization. In the overt VP configuration, it is odd, if not impossible, to raise the object of the lower VP to the subject of the matrix clause in passive constructions, as in (2.24).

(2.24) ?The paper was begun to be written by John.

If the coerced complement has the syntactic structure of the full VP configuration, it should also prohibit such object-to-subject raising. Yet passivization is well-formed for coercion expressions, as in (2.25).

(2.25) The paper was begun by John.

The third problem has to do with event anaphora. Pylkkänen (2008) indicates that the *it*-anaphor cannot refer to the silent event in sentences with complement coercion. For instance, the *it* in (2.26) cannot refer to an unstated activity involving the complement such as reading.

(2.26) #Joe enjoyed these books but it ruined his eyes.

—Examples from Pylkkänen (2008, p.719)

Pylkkänen and McElree (2006) conclude: "if coerced expressions involve a silent VP, that VP for some reason cannot be adverbially modified and it does not constitute an intervener for A-movement in passivization." Because

of these observations, they indicate that the Type-Shifting analysis for complement coercion is favored, even though it sacrifices strict compositionality.

2.3.3 Interim Summary: Linguistic analyses of complement coercion

In short, previous analyses assume that coercion verbs (e.g. *begin*, *enjoy*) require their complements to denote events and that the composition of a coercion verb with an entity-denoting complement necessarily gives rise to an eventive interpretation. The eventive interpretation is obtained by

(a) Type-Shifting: Interpolating a type-shifting operation to repair the type mismatch between the verb and the complement in semantic composition, and mining either the lexical semantics of the complement noun (Pustejovsky, 1995) or the SS-CS interface (Jackendoff, 1997, 2002) to determine the exact activity;

(b) Pragmatic inference: Inference based on contextual information to saturate an empty V head, assuming that the complement NP is syntactically identical to $[_{VP}\ [_{V^0}\ \mathbf{e}]\ [_{NP}\ ]]$ — a VP structure with an empty V (De Almeida, 2004; De Almeida and Dwivedi, 2008).

Because of the shared assumption that coercion verbs select for events as their complements, these analyses, despite their differences, yield the same pre-

diction in terms of real-time comprehension: processing the coercion configuration (i.e. an event-selecting verb + an entity-denoting complement) requires additional effort to arrive at an eventive interpretation.

2.4 The processing of complement coercion

Studies have reported that processing sentences with complement coercion is difficult during real-time comprehension. The composition of an entity-denoting complement and a coercion verb such as (2.27a) engenders additional processing cost than the composition with a non-coercion verb (2.27b) (e.g., McElree et al., 2001) or the composition of a coercion verb with an event-denoting complement (2.27c) (Traxler et al., 2002).

(2.27) a. John began/enjoyed the book.

b. John *read* the book.

c. John began the *fight*.

The complement coercion effect has been shown by a variety of methodologies[5], including eye-tracking (Frisson and McElree, 2008; Pickering et al., 2005; Traxler et al., 2002, 2005) and self-paced reading (McElree et al., 2001; Traxler

[5]However, De Almeida (2004) reports no coercion effect in a self-paced reading experiment, and suggests that the eventive interpretation in the coercion configuration results from pragmatic inference, which is cost-free.

et al., 2002), and response-signal speed-accuracy trade-off (SAT) paradigm (McElree et al., 2006). In addition, sentences with complement coercion are shown to involve specific neurological patterns (Piñango and Zurif, 2001; Pylkkänen and McElree, 2007; Pylkkänen et al., 2009; Husband et al., 2011). The question emerges is: What is the source of this processing cost? Several hypotheses have been proposed to answer this question, reviewed in the following subsections.

2.4.1 The Type-Shifting Hypothesis

The Type-Shifting analysis (discussed in Section 2.3.1) states that coercion verbs select for eventive complements, and that complement coercion involves a type-shifting operation that converts the type of the complement from an entity to an event to satisfy the selectional restriction of its governing verb. The processing cost has been hypothesized to result from the interpolation of this type-shifting operation, which is mandatory to repair the type mismatch between the event-selecting verb and the entity-denoting complement. This will be referred to as the **Type-Shifting Hypothesis**. Below I present psycho- and neuro-linguistic findings that are taken to support this hypothesis for processing complement coercion.

2.4.1.1 Psycholinguistic studies

McElree et al. (2001) report the first psycholinguistic study that evidences the complement coercion effect. In the self-paced reading experiment, they manipulate three conditions: (a) coerced, (b) preferred, and (c) non-preferred, as in (2.28). Their results show that coercion and non-preferred verb conditions engendered significantly longer reading times (RTs) than the preferred condition at the complement noun. At the complement noun+1 position, the coercion condition yielded longer RTs than both the preferred and non-preferred conditions.

(2.28) The student (a) began/ (b) read/ (c) wrote the book late in the semester.

—McElree et al. (2001)

The processing cost does not result from coercion verbs *per se* but from the composition of the verb AND its following entity-denoting complement together. This is demonstrated by Traxler et al. (2002). Their eye-tracking experiment (Exp.2) shows that, when following a same coercion verb, an entity-denoting complement induced longer RTs than an event-denoting complement in the second-pass and total time data. For example, *started the puzzle* was more taxing than *started the fight*, despite no difference in cloze probability. Their self-paced reading experiment (2002, Exp.3) reveals the same pattern.

(2.29) The boy started the fight/ puzzle after school today.

—Traxler et al. (2002)

Furthermore, when following neutral (i.e. non-coercion) verbs, entity- and event-denoting complements did not differ in cloze values (e.g. *The boy saw the fight/puzzle*). These indicate that it is the combination of [a coercion verb + an entity-denoting complement] that engenders the additional processing cost.

One might suspect that the processing cost of complement coercion could result from selecting a specific activity associated with the complement noun among a variety of possible ones. However, Frisson and McElree (2008) argue that the cost is not modulated by such competition. Their eye-tracking experiment shows that the coercion sentences with a weakly preferred reading (in which the implicit event can be interpreted as various activities) did not differ from coercion sentences with a strongly preferred reading (in which the implicit event has one dominant interpretation). For the contrasts shown in (2.30), there was no interaction between verb type (coercion vs. non-coercion) and preference (weakly vs. strongly preferred), though there was an effect of coercion at the complement NP (found in first-pass, second-pass, and total time) and its following two words (found in first-pass and total time).

(2.30) [Weakly preferred]: The teenager (a) began/ (b) read the novel...

[Strongly preferred]: The waitress (a) started/ (b) served the cof-

fee...

—Frisson & McElree (2008)

This suggests that the coercion effect does not result from the preference of sentence interpretations. The processing cost cannot be attributed to selecting an activity among several ones or ambiguity among multiple eventive readings.

In addition, Traxler, Pickering & McElree's (2005) eye-movement study (Exp.1) shows that introducing the most frequent activity in context did not attenuate the cost in the target coercion sentence.

(2.31) [Context]: The student was *reading* in his office.

[Target]: After a while, he *started a book* about health care spending.

—Traxler et al. (2005), Exp.1

Nevertheless, the cost was eliminated by a context sentence that either had the same coercion configuration (*started the book*) or explicitly mentioned the entire event structure (*read the book*). While the coercion effect was observed in the context sentence alone (2.32a > 2.32b), such effect was absent in the target sentences following the context (2.33a = 2.33b).

(2.32) [Context]: The student (a) *started* /(b) *read* a book in his dorm room.

(2.33) [Target]: Before he (a) *started* /(b) *read* a book about opium trade, he checked his e-mail.

—Traxler et al. (2005), Exp.3

Moreover, the coercion effect is attained in the contrast between coercion sentences (coercion verb + entity NP) and sentences containing the full-VP (coercion verb + V-ing + entity NP). In Pickering, McElree, & Traxler's (2005) eye-movement study, the coercion condition (2.34a) induced longer RTs than any of the conditions in (2.34b), including the preferred reading (*built*), non-preferred reading (*sanded*), and full-VP conditions (*began building*) in total time data at the complement NP region (*the table*). The three conditions in (2.34b) did not differ from one another. At the post-noun region (*during the*), the coercion condition also induced more first-pass regressions than the full-VP condition.

(2.34) a. The carpenter began the table during the morning break.

b. The carpenter built/ sanded/ began building the table during the morning break.

—Pickering, McElree, & Traxler (2005)

In sum, results from these psycholinguistic studies suggest that the difficulty of processing complement coercion is engendered by the composition of the assumed event-selecting verbs and their following entity-denoting complement NP. It is held that the cost results from type-shifting the entity-denoting complement to build an event representation, to fulfill the selectional restriction of the event-selecting coercion verbs. Importantly, the complement coercion effect cannot be attributed to the verb *per se* (Traxler et al., 2002),

the selection of a specific activity among multiple plausible ones by preference (Frisson and McElree, 2008), the retrieval of a suitable activity in context (Traxler et al., 2005), or the co-occurrence frequency of the V-N combination (Pylkkänen and McElree, 2006).

2.4.1.2 Neurolinguistic studies

Kuperberg et al. (2010b) conducted an ERP study contrasting three conditions: (a) coercion, (b) animacy violation, and (c) non-coercion control, as in (2.35).

(2.35) The journalist (a) began/ (b) astonished (c) wrote the article before his coffee break.

—Kuperberg et al. (2010)

They found that complement coercion and animacy violation induced larger N400 than the control, with similar widespread scalp distribution and amplitude at the complement noun. On the other hand, animacy violation induced a prolonged P600 effect in the posterior region, relative to both coercion and control at the same window. Noteworthily, at the sentence-final word, coercion sentences evoked a prolonged anterior positivity as compared to the other two conditions. Kuperberg et al. suggest that the N400 reflects the mismatch between the semantics of the verb and the complement, while the P600 reflects the effort of retrieving the unstated activity as a result of the verb-complement

composition.

Using the same contrast (coerced, animacy-anomalous, and non-coerced), Baggio et al. (2010) report that complement coercion engendered an N400 effect from the complement noun onset, and this negativity lasted until the end of the epoch (1000ms) with a central scalp distribution. The animacy-anomalous sentences also induced an N400 effect, with similar scalp distribution as the coerced sentences, but this negativity did not last long. That is, the coerced and anomalous conditions differed not in the cortical distribution but in *duration*. They suggest that the N400 and the following sustained negativity reflect an inference-driven unification that integrates the lexical items into a complex semantic representation (one with the missing activity) and inferences the suitable activity in the given context.

Notice that the two ERP studies differ in that sentences involving coercion engendered a prolonged *positivity* following N400 in Kuperberg et al. (2010b) but a prolonged *negativity* following N400 in Baggio et al. (2010). This discrepancy needs to be clarified by future studies.

In terms of neural correlates associated with processing sentences involving complement coercion, three distinct brain regions have been reported. Piñango and Zurif's (2001) lesion study contrasts transparent non-coercion sentences (a) versus their coercion counterparts (b), as in (2.36).

(2.36) a. The boy *began reading* the book.

b. The boy *began* the book.

They found that Wernicke's aphasics had difficulty comprehending sentences with complement coercion in a sentence-picture matching task, whereas Broca's aphasics and the control group did not show difficulty. Thus, the authors suggest that Wernicke's area supports the type-shifting operation which introduces the event in semantic composition.

However, a different region is reported in Pylkkänen & McElree's (2007) MEG study. They contrast coerced (*began*), anomalous (*astonished*), and control (*wrote*) sentences as in (2.37).

(2.37) The journalist began/ astonished/ wrote the article after his coffee break.

—Pylkkänen and McElree (2007)

At the complement noun, the coerced condition, as compared to both the anomalous and control conditions, increased MEG signals in the anterior midline field (AMF), which was generated in ventrol-medial prefrontal cortex (vmPFC) at 350-450ms. In contrast, semantically anomalous sentences induced a left M350 but no AMF activity. The authors indicate that the AMF-vmPFC effect associated with complement coercion results from construing an implicit activity from the complement noun as requested by the event-selecting verbs.

On the other hand, Husband et al.'s (2011) fMRI study reports yet another region associated with complement coercion. Contrasting coercion (*began*),

semantic violation (*annoyed*), syntactic violation (*write*), and control (*wrote*) sentences as in (2.38), they found that coercion sentences engendered additional activity in left inferior frontal gyrus (LIFG), in BA45 more specifically, compared to control. This LIFG activity is attributed to the detection of a semantic type mismatch and the type-shifting operation.

(2.38) The novelist began/ annoyed/ write/ wrote the book before break.

—Husband et al. (2011)

In short, neurolinguistic experiments investigating complement coercion show discrepant results: Wernicke's area (Piñango and Zurif, 2001), vmPFC (Pylkkänen and McElree, 2007), and LIFG (Husband et al., 2011). The inconsistency remains to be explained by further studies. Lai et al.'s (2014) fMRI experiment found both Wernicke's area and LIFG for sentences like "*John began the book*," but the hypothesis suggested there is different from the Type-Shifting hypothesis. This study will be presented in Chapter 4.

2.4.2 Challenges to the Type-Shifting Hypothesis

Recall the fundamental assumption of the Type-Shifting Hypothesis: coercion verbs like *begin* and *enjoy* impose a selectional restriction on their complement, requiring it to denote an event. Under this assumption, sentences

with coercion verbs always yield an eventive interpretation. However, the Type-Shifting Hypothesis encounters at least two challenges, pointed out in Piñango & Deo, (2012, 2015).

First, in terms of linguistic distribution, the composition of a coercion verb with an entity-denoting complement does not necessarily give rise to an eventive interpretation. Consider the sentences in (2.39) from the Corpus of Contemporary American English (COCA):

(2.39) a. This image begins the genealogy of the kings of England and flows into materials specifically written for St. Albans

b. "In the beginning was the Word" begins the prologue of John's gospel (1:1).

c. The sun begins the day, darkness ends it.

d. Silence finishes the conversation.

e. The house is outlined in white lights, with candles and stars in the windows. A moving train finished the display.

f. Although this is mostly a collection of previously published essays, it is notable because of the new autobiographical memoir that begins the book.

g. This column continues the review of the new PSA Club Services Website.

Sentences in (2.39) all contain a coercion verb followed by an entity-denoting

complement, yet the complement is not construed as an event. In other words, there is no "complement coercion" in these examples. This casts doubt on the assumption that coercion verbs select for eventive complements.

Second, in terms of real-time comprehension, it turns out that only a subset of coercion verbs engender additional processing cost. Katsika et al.'s (2012) eye-movement study contrasts aspectual verbs, which denote temporal reference such as *begin* and *finish*, with psychological verbs, which denote mental states such as *enjoy* and *endure*. Their results show that the processing cost is observed only for aspectual verbs, but not psychological verbs.

Utt et al.'s (2013) study also suggests a dissociation between aspectual and psychological verbs. Using a corpus-based approach, they measure "eventhood that captures the degree to which verbs expect objects that are events rather than entities." Eventhood is defined as the percentage of the complement nouns that have an event sense. If coercion verbs[6] do select for events, these verbs should be associated with eventive complements statistically more and hence yield higher eventhood scores than non-coercion verbs. They test two verb sets: (a) those in Traxler et al. (2002), in which verbs are dichotomized as coercion and non-coercion verbs, and (b) those in Katsika et al. (2012), in which verbs are categorized as aspectual, psychological, and non-coercion verbs. Results show that, with Katsika's stimuli, aspectual verbs yielded significantly higher eventhood scores than psychological verbs. In contrast, Traxler,

[6]In Utt et al. (2013), coercion verbs are referred to as "metonymic verbs."

Pickering, & McElree's stimuli, which collapse aspectual and psychological verbs, did not yield significance in eventhood scores. In addition, when non-aspectual verbs (*enjoy, endure, expect, prefer*) were excluded from the set of coercion verbs, the difference between coercion verbs (now aspectual verbs only) and non-coercion verbs became significant. The pattern supports the dissociation between aspectual and psychological verbs as made in Katsika et al. (2012).

In closer examination, the set of "coercion verbs" used in previous studies include aspectual verbs (e.g. *begin*), psychological verbs (e.g. *enjoy*), and some control verbs (e.g. *try, master*). Katsika et al.'s (2012) and Utt et al.'s (2013) findings suggest that the set of coercion verbs in fact conflate heterogeneous verbs classes. These verb classes should not receive the same analysis as Type-Shifting, given that subclasses within the coercion verb set do not exhibit an identical processing profile—the cost is associated with only aspectual verbs but not others. This means that previous experiments on complement coercion are subject to the problem of heterogeneous stimuli.

The above discussion presents two challenges for the Type-Shifting Hypothesis. In terms of empirical coverage, not all sentences with coercion verbs give rise to eventive interpretations. In terms of experimental implementation, the study of Katsika et al. (2012) shows that only aspectual verbs, but not psychological verbs, induce additional processing cost. These problems call for a need to re-examine the different verb types collapsed as the set coercion

verbs, as well as the processing of complement coercion.

2.4.2.1 Interim Summary: the Type-Shifting Hypothesis

The experimental work that assumes the assumption of the Type-Shifting Hypothesis, i.e., coercion verbs select for events, is summarized in Table 2.1. These studies attribute the additional processing cost and brain activity to a process that generates the eventive interpretation for sentences with complement coercion. Nevertheless, Section 2.4.2 points out that the Type-Shifting Hypothesis is problematic in terms of linguistic distribution and psychological implementation. A new analysis which singles out aspectual verbs from others is needed to capture the extra effort of processing sentences with these verbs. Chapter 4 will present such an alternative account for aspectual verbs.

Table 2.1: Experiments on complement coercion

Study	**Conditions**	**Findings**
McElree et al. (2001) Self-paced reading	[Coerced/Preferred/Nonpreferred] The student *began/read/wrote* the book late in the semester.	Coercion engendered longer RTs at complement noun and the following word.
Traxler et al. (2002) Eye-movements & Self-reading	**Exp.1** The secretary *began/typed/read* the memo about the new office policy. **Exp.2 & 3** The boy started the *fight/puzzle* after school today.	Coercion increased RTs at the complement and its following word in first-pass regressions and total time.
Lapata et al. (2003) Sentence completion in German	Subjects that trigger [Telic / Agentive / Neutral] qualia roles X verb X object The *student/author/Peter* [began/regretted/enjoyed] the [book/speech/...] ___ (V) ___.	(a) Neutral and Telic subjects biased towards Telic completion; Agentive subjects biased towards Agentive completion. (b) Different [V+Obj] combinations yield distinct reading preference, regardless of subject.
De Almeida (2004) Self-paced reading	**Exp.1** [Coerced/Preferred/Nonpreferred] Materials modified from McElree et al. (2001)	No coercion effect was found at the complement noun and N+1.
	Exp.2 Context + Target (Context): The student was a procrastinator.... (Target): The student *begin/read/wrote* the paper...	At complement N and N+1, coerced and nonpreferred conditions induced longer RTs than preferred.
Pickering et al. (2005) Eye-movement	[Coerced/Preferred/Nonpreferred/Ude VP] The carpenter *began/ built/ sanded/ began building* the table.....	Coercion showed greater RTs at the noun and post-noun regions, compared to all others, in first-pass regressions and total time.
Traxler et al. (2005) Eye-movement	**Exp.1** (Context): The student was *reading/resting* in his office. (Target): .....he *started/read* a book about...	Coercion induced longer total fixation time regardless of context.

	Exp.3 (Context): The student *started/read* a book in his dorm room. (Target): Before he *started/read* the book about...	Coercion cost was eliminated by context that has the same coercion configuration or provides the entire event structure.
	Exp.4 Same contrast but the complement NP was replaced by pronouns.	Same pattern as in Exp.3
McElree et al. (2006) Multiresponse speed-accuracy trade-off paradigm (MR-SAT)	[Coerced/Noncoerced] The carpenter *began/built* the table.	Coerced sentences are less likely to give rise to an acceptable interpretation and readers spent more time computing an acceptable reading.
Frisson and McElree (2008) Eye-movements	[Coercion/Control] a. (Strongly preferred, with a dominate reading): The teenager *began/read* the novel... b. (Weakly preferred, without a dominate reading): The waitress *started/served* the coffee...	Coercion induced longer RTs at the complement region in first-pass, second-pass, and total-time. The coercion effect also appeared at the complement+1 region, in first-pass and total time data. No effect of preference was found.
Kuperberg et al. (2010b) ERP	[Coerced/Animacy-violated/ Control] The journalist *began/ astonished/ wrote* the article...	Coercion evoked an N400 effect and a sustained anteriorly distributed positivity at sentence-final words.
Baggio et al. (2010) ERP	[Coerced/Anomalous/Control] The journalist *began/ astonished/ wrote* the article...	Coercion evoked an N400 effect and a prolonged negativity from 700ms∼1000ms with a central scalp distribution.
Piñango and Zurif (2001) Lesion study	[Transparent/Coercion] The boy *began reading/ began* the book.	Wernicke's aphasics performed worse than both Broca's aphasics and control group on coercion sentences.

Pylkkänen and McElree (2007) MEG	[Coerced/Anomalous/Control] The journalist *began/ astonished/ wrote* the article after his coffee break.	Coercion engendered signals in midline frontal field (AMF) generated from vmPFC at 350-450ms at the complement noun.
Pylkkänen et al. (2009) MEG	[Coercion(N_{entity})/ Control(N_{event})] The nimble climber imaged the *ice/fall* survivable...	Coercion triggered an AMF effect at 650-700ms at the able-suffixed adjectives.
Husband et al. (2011) fMRI	[Coerced/ Semantic violation/ Syntactic violation/ Control] The novelist *began/ annoyed/ write/ wrote* the book before the break	Coercion recruited LIFG (BA45) as compared to control.

2.4.3 The Surprisal Hypothesis

2.4.3.1 The Surprisal Hypothesis

Recently, Delogu and colleagues (Delogu and Crocker, 2012; Delogu et al., 2013, 2017) argue that the difficulty with coercion sentences could arise from the predictability of the complement noun, which tends to be lower in the coercion configuration than in the control configuration. The lower predictability of the complement noun in coercion sentences, as compared to the non-coercion sentences, results in *surprisal* during sentence comprehension and therefore the cost. Delogu et al. (2017) note that the observed cost associated with coercion "is not a consequence of recovering enriched representation *per se*, but rather their lower predictiveness prior to the coerced noun, and thus higher

surprisal."

Delogu and Crocker (2012) present an eye-tracking experiment in German, contrasting three verb types— coerced (*began*), preferred (*read*), neutral (*bought*)— as shown in (2.40).

(2.40) Peter began/ read/ bought the book [on holiday].

The Surprisal Hypothesis predicts that the preferred condition, a more predictive configuration, will be easier to process than the other two conditions. On the other hand, the Type-Shifting Hypothesis predicts that both the neutral and preferred conditions will be easier than the coerced condition because only the latter triggers the type-shifting operation.

(A) The Surprisal Hypothesis	[Coerced=Neutral] >[Preferred]
(B) The Type-Shifting Hypothesis	Coerced >[Neutral=Preferred]

Results show the coerced condition engendered longer RTs than both the preferred and neutral conditions in the complement noun region only in total time data, consistent with the Type-Shifting Hypothesis. However, in the spill-over region (square-bracketed in (2.40)), both the coerced and neutral conditions engendered comparably longer RTs than the preferred condition in first-pass regressions and regression-path times, consistent with the Surprisal Hypothesis. In addition, in an ERP experiment, Delogu et al. (2017) found a distributed N400 effect for both the coerced and neutral conditions relative to

the preferred condition.

According to these authors, the above findings suggest that the coercion cost is initially driven by surprisal for the less predicted noun (indicated by N400 and early effect in the spillover region) and subsequently by the type-shifting operation in a later stage (indicated by the total time data on the complement noun). To the extent that Delogu et al. (2017) still attribute the cost unexplained by predictability to the coercion operation, they retain the assumption that coercion verbs select for eventive complements. That is, what they suggest is that "the coercion cost largely reflects the surprisal of the complement noun" but part of the cost still results from the type-shifting operation.

2.4.3.2 Challenges to the Surprisal Hypothesis

The Surprisal Hypothesis attributes the coercion cost to the less predictive (less probable) Verb-Object combination in the coercion configuration (i.e. event-selecting verb + entity-denoting noun). Pylkkänen and McElree (2006), using Traxler, Pickering, & McElree.'s (2002) stimuli, examine the co-occurrence of event- vs. entity-denoting complements following a coercion or non-coercion verb. As it turns out, event-denoting NPs co-occur *less* with either coercion verbs or neutral verbs than entity-denoting NPs . In other words, the [coercion verb + entity-denoting complement] combination requires more

processing effort even though it is more frequent. Hence, the online cost cannot be reduced to the distributional likelihood of the coercion configuration.

The Surprisal Hypothesis is further questioned by DiNardo's (2015) Cloze study. She contrasts three verb types: aspectual verbs (e.g. *begin, finish*), psychological verb (e.g. *enjoy, endure*), and non-coercion psychological verbs such as *love, hate* as control. Notice that in previous studies, the former two types are considered as coercion verbs whereas the latter is not. In a Verb Multiple Choice experiment, the participants were given sentences with an animate subject referent and an entity-denoting complement like (2.41) and asked to rate the sentences with different verb types from a 1∼5 scale. Results show that aspectual and psychological verbs—the coercion verb set—were rated above 3.9, suggesting that such construction is not dis-preferred.

(2.41) F. Scott Fitzgerald _____ the anthology of 20th century American literature.

Also, in the Fill-in-the-Blank experiment, the participants were asked to write the first two verbs that came to their mind for the sentences. Results show that three coercion verbs (*began, started, enjoyed*) were among the top-ten verbs chosen by the participant.[7] Although the most frequently chosen ones were action verbs such as *wrote* and *read*, it is likely a result of Gricean

[7]The ten verbs most used by the participants in the Fill-in-the-Blank experiment include (with decreasing counts of usage noted in parentheses): *wrote* (256), *read* (129), *loved* (50), *hated* (50), *enjoyed* (42), *watched* (40), *compiled* (38), *published* (34), *started* (32), *began* (30) (DiNardo, 2015).

maxim of Manner, which urges the participants to be clear in expressing meaning and avoid ambiguity.

Besides, the Surprisal Hypothesis is subject to the problem of heterogeneous stimuli as well, just as the Type-Shifting Hypothesis is. The experiments presented in Delogu et al. (2012, 2013, 2017) conflate aspectual and psychological verbs as the set of coercion verbs, yet the two verb types have been shown to display different processing profiles (Katsika et al., 2012).

The findings of Pylkkänen and McElree (2006) and DiNardo's (2015) suggest that the coercion configuration is not unexpected to comprehenders, and therefore the processing cost is unlikely to result from surprisal. In addition, as Delogu et al.'s (2012, 2013) admit, their findings also support the Type-Shifting Hypothesis; the Surprisal Hypothesis does not appear to win out.

2.5 Conclusion

In this chapter, I discussed the complement coercion phenomenon, which manifests enriched composition and semantics-syntax mismatch. To recapitulate, it refers to the observation that the composition of a coercion verb (i.e. verbs assumed to semantically select for an eventive complement) and an entity-denoting complement often gives rise to an eventive interpretation. I reviewed the analyses that attempt to account for the eventive interpretation. On the Type-Shifting analysis, the eventive interpretation is generated by a

type-shifting operation that coerces the semantic type of the complement from an entity to an event, in order to fulfill the selectional restriction of coercion verbs. In Pustejovsky's (1991, 1995) Generative Lexicon framework, the specific event is obtained via the qualia structure of the complement noun. In Jackendoff's (1997) Parallel Architecture, the unexpressed event is obtained by interpolating a reinterpreting function in the SS-CS interface. In contrast, de Almeida (2004) and de Almeida & Dwivedi (2008) claim that the eventive interpretation is computed by pragmatic inference to realize the missing V-head in the syntactic structure, assuming that the complement of coercion verbs is underlyingly a VP with an empty V head.

Experimental work shows that sentences with complement coercion engender additional processing cost during real-time comprehension. The Type-Shifting Hypothesis states that the cost results from the insertion of the type-shifting operation to repair the mismatch between the verb and the complement, construing the complement as an event. On the other hand, the Surprisal Hypothesis claims that the cost results from the less predictive $\text{Verb}_{coercion}$–Noun_{entity} combination.

It is important to point out that these analyses and hypotheses are based on the same assumption: coercion verbs (e.g. *begin, enjoy, master*) require an event-denoting complement, and coercion sentences necessarily give rise to an eventive interpretation (e.g. "*begin/enjoy/master the book*" is read as "*begin/enjoy/master V-ing the book*"). Nevertheless, this assumption is chal-

lenged by empirical observations, and the set of coercion verbs is shown heterogeneous, exhibiting different processing profiles.

In the next chapter, I will introduce a semantic analysis by Piñango and Deo (2012, 2015). It targets a subset of coercion verbs, namely aspectual verbs based on their semantic properties. Their analysis leads to the hypothesis that captures the cost associated with complement coercion as the processing of aspectual verbs.

Chapter 3

The Structured Individual Hypothesis

3.1 Introduction

In Chapter 2, I presented different accounts that attempt to capture the complement coercion phenomenon. They all assume that coercion verbs—including aspectual verbs (e.g. *begin*) and psychological verbs (e.g., *enjoy*) place a selectional restriction on their complement. However, as pointed out in Section 2.4.2 of Chapter 2, such assumption does not hold in example (2.39), where a coercion verb is combined with an entity-denoting complement yet no eventive reading is engendered. In addition, the study by Katsika et al. (2012) suggests that the processing cost is observed only for aspectual verbs (e.g., *begin, finish*)

but not psychological verbs (e.g., *enjoy, prefer*). Based on these observations, several questions emerge:

(i) How should we explain the phenomenon that sentences with verbs like *start, finish* (identified as coercion verbs in previous accounts) give rise to both an eventive reading as in "*John began the book*" and a non-eventive reading as in "*Silence finishes the conversation.* (=2.39d)"?

(ii) If an entity-denoting complement is not coerced into an event in sentences with a non-eventive reading like (2.39), how should we account for the processing cost previous attributed to complement coercion?

In what follows, I introduce Piñango & Deo's (2012, 2015) Structured Individual analysis (Section 3.2) that aims to answer question (i). This analysis paves the way for the Structured Individual Hypothesis (Section 3.3) that provides an answer to question (ii).

3.2 The Structured Individual analysis (Piñango & Deo 2012, 2015)

3.2.1 Lexical semantics of aspectual verbs

Piñango & Deo's (2012, 2015) analysis provides a unified analysis for aspectual verbs such as *begin, continue, end*, which appear in sentences that

yield an eventive reading "*John began the book*" (the coercion configuration) and sentences that yield a non-eventive reading as in (2.39). They argue that aspectual verbs (AspVs) lexically select for complements that denote *structured individuals*, rather than events. Structured individuals are defined as entities that can be mapped onto a one-dimensional directed path structure (DPS) along some ontological dimensions (e.g. temporal, spatial, eventive, etc.). A DPS is defined, following Krifka (1998), as a totally ordered structure whose adjacent (non-overlapping) parts exhibit a precedence relation along some dimensions.

(3.1) *Directed Path Structure*

For example, a three-dimensional spatial entity like "a tunnel" can be mapped onto a line (a one-dimensional DPS) along its most salient dimension, typically spatial (Jackendoff, 1992; Verkuyl and Zwarts, 1992). Likewise, "a book" may be construed as a one-dimensional path structure along the informational dimension, taking its chapters as the adjacent parts. Below are more examples from the Corpus of Contemporary American English (COCA) that further illustrate the use of aspectual verbs.

(3.2) a. A moving train finishes the display. (Spatial)

b. This January 1 begins the dawn of a new age of attainable resolution. (Temporal)

c. His death begins the Revel. (Eventive)

d. The only adequate or appropriate response to this reality seems to be the expression that both begins and ends the novel. (Informational)

An entity that is a one-dimensional DPS is called an "axis" by Piñango & Deo (p.20). The notion of structured individuals makes reference to axes onto which individuals are mapped. More specifically, structured individuals are defined as entities that are relatable to such axes via homomorphic functions. An individual x of any type τ is taken to be a structured individual relative to a function f of any type (τ, σ) iff $f(x)$ is an axis and f is a homomorphism from the part structure of x to the axis $f(x)$.

(3.3) $\forall x_\tau[\textbf{struct-ind}_{f_{<\tau,\sigma>}}(x) \leftrightarrow [\textbf{axis}(f(x)) \wedge \forall x', x'' \leq x[x' \leq x'' \rightarrow f(x') \leq f(x'')]]]$

—Piñango and Deo (2015, p.21)

Aspectual verbs carry a lexical presupposition that requires their complement to denote a structured individual. Truth-conditionally, they require the subject to be mapped onto some privileged small subpart of the axis determined by the complement denotation. A sentence with an aspectual verb is true iff the subject denotation is construed as a specific type of (e.g. initial, medial, final) subpart of the axis that the complement denotation is mapped to.

Moreover, aspectual verbs lexically encode a set of dimension functions (f_c) that are contextually accessed; each function maps the complement denotation to an axis along a specific dimension (e.g. f_{space}, f_{time}, f_{info}). Aspectual verbs are sensitive to whether their complement can be construed in context as a structured individual, i.e. whether it is possible to map it to an axis along some dimension in context. The contextually accessed function f_c reflects this contextual reliance. Once the presupposition that the complement denotes a structured individual is met, aspectual verbs require that there be some function such that it maps the subject denotation to a small initial part ($<_{small-init}$) of the axis given by f_c(x). (3.4a) provides a sample lexical entry for "*begin*."

(3.4) a. $[\![\text{begin}]\!] = \lambda x_\tau \lambda y_\sigma : \textbf{struct-ind}_{f_c}(x).\exists f'[f'(y) <_{small-init} f_c(x)]$

b. *Begin(x)(y)* is defined iff x is a structured individual with respect to the contextually determined function f_c. If defined, *begin(x)(y)* is true iff there is some function f' (possibly identical to f) such that $f'(y)$ is a small initial subpart of the axis $f_c(x)$.

—Piñango and Deo (2015, p.21)

By making reference to structured individuals as axes, this proposal generalizes the lexical meaning of aspectual verbs to accommodate arguments of different semantic types, accounting for a broader range of distribution.

In addition, the lexical meaning of the complement must encode informa-

tion about the possible dimensions regarding the concept denoted by the complement. For instance, an expression like "*book*" must contain the information about its spatial as well as informational structure; expressions like "*path*" and "*river*" must contain the information that they have a salient spatial dimension. This information is mined during the composition with aspectual verbs, for determining the dimension along which the complement denotation is to be construed as a structured individual in a given context. Crucially, the selection of the functions that map the complement and subject denotations to an axis along a specific dimension depends on both the properties of the complement and context.

3.2.2 The notion of *Dimension*

On the Structured Individual analysis, the reading of a sentence with an aspectual verb in underdetermined. The full interpretation depends on the specific dimension for structuredness chosen in a given context. Let me explicate the notion of dimension with examples to illustrate.

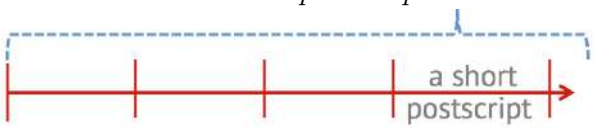
In "*This famous perch begins the Appalachian Trial*," the complement (*the Appalachian Trail*) is conceptualized as a structured individual, construed as a DPS/axis along the **spatial dimension**. The spatial function (f_{space}) encoded as part of lexical content of the verb maps the subject denotation to the initial subpart of the axis construed by the complement. The sentence is true iff the

spatial extent of *this famous perch* is the initial subpart of the spatial extent of *the Appalachian Trail* (3.5).

(3.5) SPATIAL: *This famous perch begins the Appalachian Trail.*

In "*A short postscript ends the novel*," the complement is conceptualized as a structured individual construed as an axis along the **informational dimension**. The informational function (f_{info}) of the verb "*end*" maps the subject denotation onto the final subpart of this axis. The sentence is true iff the informational content denoted by *a short postscript* is the final subpart of the set of informational content denoted by *the novel* (3.6).

(3.6) INFORMATIONAL: *A short postscript ends the novel.*



In "*A prayer started the banquet*," the axis construed by the complement is realized along the **eventive dimension**. In this case, the verb maps the subject denotation onto the initial subpart of the axis. The sentence is true iff the event denoted by *a prayer* is the initial sub-event of the event denoted by *the banquet* (3.7).

(3.7) EVENTIVE: *A prayer started the banquet.*

These examples illustrate that a structured individual can be construed as a one-dimensional DPS (i.e., an axis) in a variety of dimensions. Comprehenders have to rely on the context to determine the exact dimension along which the structured individual is construed as an axis. For instance, in the context of formatting or printing a book, the spatial dimension will be chosen; in the context of reading a book, the informational dimension is more relevant.

3.2.3 The coercion configuration of aspectual verbs

How does this Structured Individual analysis account for the eventive interpretations of the "coercion" configuration—animate subject referent + AspV + entity-denoting complement—such that (3.8a) is interpreted as (3.8b)?

(3.8) a. William Shakespeare began the book.

b. ⇝ William Shakespeare began *reading/writing/...* the book.

In the coercion configuration, sentences receive a reading in which the animate subject referent is the agent of some implicit eventuality. Piñango & Deo analyze these cases as involving functions that map entities to the events which they are participants of; these are termed "inverse thematic functions." While thematic roles (e.g. agent, patient, theme) map events to their participants as the actor or undergoer, inverse thematic functions are defined as functions that

"map pairs of individuals and times to the smallest event that the individual bears a participant role to at that time in a given context" (Piñango and Deo, 2015, p.27). In other words, inverse thematic functions are ways of accessing events via the individuals that participate in the events, rather than accessing individuals via the participant roles that they bear in an event. For instance, the inverse thematic function f_{ag_i} maps an individual to an event that it is an agentive participant of at the reference interval i in a given context. Likewise, f_{th_i} maps an individual to the event that it is the patient/theme of at i in a given context.

The reading (3.8b) for (3.8a) is an **agentive reading** along the eventive dimension. Here the axis is construed as an event of which *William Shakespeare* and *the book* are participants. The function f_c is taken to be f_{th_i}, which maps the complement (*the book*, conceptualized as a structured individual) to the smallest event of which it is the theme at the reference time. The subject (*William Shakespeare*) is mapped to the smallest event of which it is the agent by the inverse thematic function f_{ag_i}. The sentence is true iff the event of which *William Shakespeare* is an agentive participant at a reference interval i in a given context is an initial part of the event of which *the book* is a theme participant at i.

Notice that the coercion configuration like (3.8a) also gives rise to a **constitutive reading**, in which the individuals—William Shakespeare & the book—get mapped to an axis along the informational (or other) dimension.

In this case, the informational content corresponding to *William Shakespeare* is understood to be a small initial subpart of the informational content corresponding to *the book*. Such an interpretation can be paraphrased as "[*The work of William Shakespeare*] *began the book*," which involves metonymy applied to the subject. The subject denotation is mapped to a subpart of the axis construed by the complement denotation.

To emphasize, on the Structured Individual analysis, the coercion configuration of aspectual verbs is underspecified with respect to the agentive and the constitutive readings (though one reading might be more salient than the other in a particular context). Both are theoretically plausible and will be disambiguated when enough contextual information is supplied. The complement can be construed as an axis along various dimensions, and the sentence containing an aspectual verb is interpreted accordingly.

In sum, the Structured Individual analysis suggests that aspectual verbs require its complement to denote a structured individual that can be construed as an axis (one-dimensional DPS) along a variety of dimensions (e.g. spatial, temporal, informational, eventive). The full interpretation of sentences with aspectual verbs requires the determination of the appropriate dimension along which the structured individual denoted by the complement is mapped onto. Unlike the Type-Shifting analysis, the Structured Individual analysis argues that the composition of an aspectual verb and an entity-denoting complement involves no semantic mismatch, and thus the type-shifting operation is unnec-

essary. In other words, there is no complement coercion at all. The advantage of this analysis is that it accounts for both the agentive and the constitutive readings associated with aspectual verbs in a unified fashion. The constitutive reading is not taken into consideration, and is left unexplained, by the Type-Shifting and the Surprisal accounts mentioned in Chapter 2.

3.3 The Structured Individual Hypothesis and predictions

3.3.1 The Structured Individual Hypothesis

Now let me address the question: What is the source of the processing cost associated with sentences containing aspectual verbs, as in "*John began the book*"? This question is important because pinning down the source of the cost adjudicates between different analyses and hypotheses of the complement coercion phenomenon.

The Structured Individual analysis for aspectual verbs leads to the Structured Individual Hypothesis (SIH) that captures the processing cost associated with these verbs. Following Piñango & Deo, this hypothesis takes aspectual verbs to encode a set of dimension functions (e.g. f_{space}, f_{time}, f_{info}) stored in their lexical entries. Aspectual verbs require their complement to be concep-

tualized as a structured individual, which is mapped to an axis along some dimensions. Sentences with aspectual verbs are underspecified among multiple dimension readings, and the exact interpretation hinges on which dimension is chosen in a given context.

Specifically, the real-time comprehension of a sentence with an aspectual verb proceeds incrementally as below (3.9). The corresponding subprocesses are presented in Figure 3.1.

(3.9) (A) **Lexical retrieval**: Upon encountering an aspectual verb, comprehenders exhaustively retrieve all lexical functions encoded in the verb. Each function is associated with a specific dimension, mapping the individuals denoted by the subject onto a subpart of the axis construed from one of the dimension associated with the complement denotation.

(B) **Resolution of dimension ambiguity**: When comprehenders encounter the complement, the complement is conceived as a structured individual as required by the aspectual verb. Because the structured individual can be construed as an axis along a variety of dimensions, the composition of the verb and the complement gives rise to multiple dimension representations, resulting in a dimension ambiguity. Comprehenders have to determine in context the relevant dimension along which the structured individual

is construed, in order to obtain an appropriate interpretation for the sentence.

The Structured Individual Hypothesis attributes the observed cost to (A) the exhaustive retrieval of dimension-functions in aspectual verbs, and (B) the resolution of dimension ambiguity (i.e., determining the contextually-relevant dimension) after the composition with the complement. The two processes have been independently motivated by Shapiro et al. (1989)[1] and Frazier and Rayner (1990)[2].

Notice that while dimension ambiguity is resolved by contextual constraints, the dimension composition of the verb and the complement (Figure 3.2) takes place regardless of context. On the SIH, the role of context is to constrain (i.e., narrow down) the dimension readings after multiple dimension representations are composed.

The SIH is distinct from the Type-Shifting Hypothesis, which attributes the processing cost to the interpolation of an operation that shifts the type of the entity-denoting complement to an event (e.g., McElree et al., 2001; Traxler et al., 2002, 2005; Pylkkänen, 2008). The SIH also differs from the Surprisal Hypothesis (Delogu and Crocker, 2012; Delogu et al., 2017), which

[1]Shapiro et al. (1989) showed that verbs with more possible argument structure arrangements (dative verbs) increased response time to the secondary lexical decision task than those with fewer possible arrangements (transitive verbs), regardless of the context. The authors therefore suggest that all the argument structure information of a verb must be momentarily activated during real-time processing.

[2]Frazier and Rayner (1990) found that delaying the disambiguation information for sentences that contain a word with multiple meanings increased processing cost.

Figure 3.1: The Structured Individual Hypothesis for processing aspectual verbs

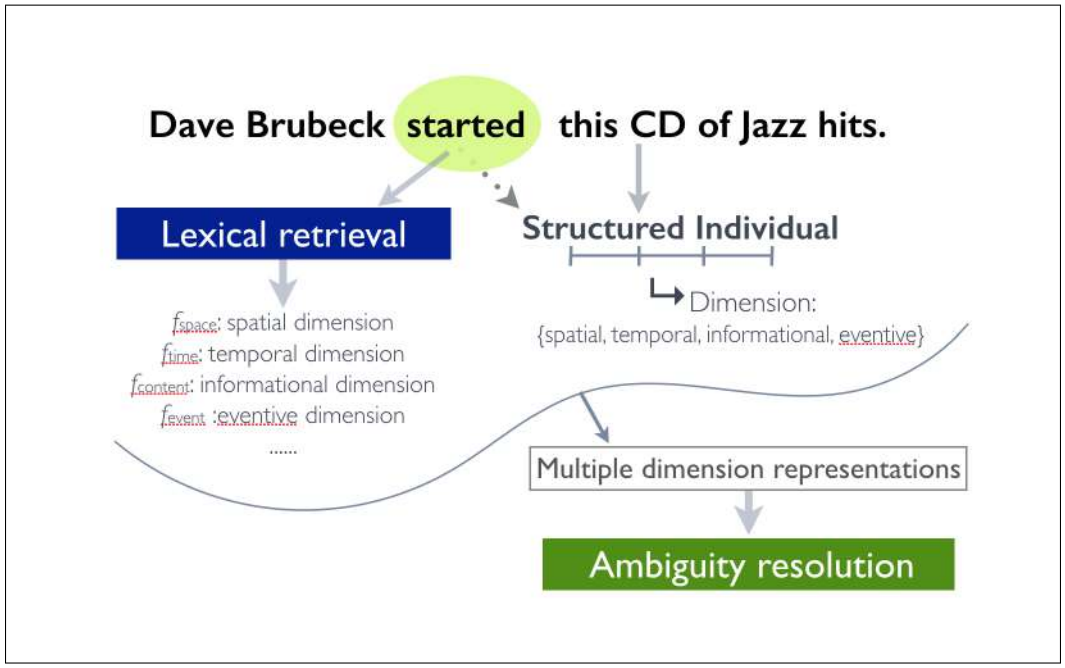
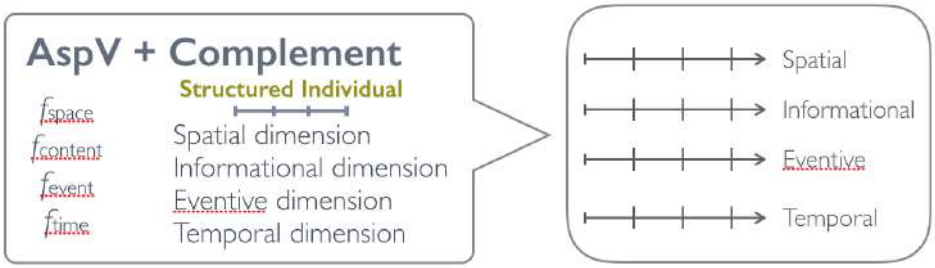


Figure 3.2: The dimension composition of an aspectual verb and its complement



claims that the cost is associated with the predictability of the complement noun and subsequent type-shifting operation. In addition, the SIH is also in contrast to the pragmatic inference account, which claims that the eventive reading is obtained via pragmatic inference and hence does not engender cost (De Almeida, 2004; De Almeida and Dwivedi, 2008).

3.3.2 Predictions of the Structured Individual Hypothesis

The Structured Individual Hypothesis (SIH) yields two sets of predictions regarding the real-time comprehension of aspectual verbs.

First, processing sentences with aspectual verbs will require additional processing effort because it involves the two processes, (A) and (B), mentioned above. Notice that the SIH dissociates aspectual verbs from other coercion verbs such as psychological verbs (e.g., *enjoy, endure*) and others (e.g., *try, master*). The SIH predicts that aspectual verbs, due to their semantic properties, will engender distinct behavioral and neurological patterns as compared to these other verb classes.

The dissociation of the verb types already gains support from a few studies. Katsika et al. (2012) found that sentences with aspectual verbs engendered more processing cost than sentences with psychological verbs. The study of Utt et al. (2013) also support the dissociation of the two verb types, based on the

differentiated co-occurrence with eventive nominals. They show that aspectual verbs co-occur significantly more with eventive nominals that denotes actions, cognitive processes, or biological processes, such as "*fight*," than psychological verbs. While this could be compatible with the Type-Shifting account, the distinction within the set of coercion verbs appears rather surprising. It is left unexplained why aspectual verbs occur more frequently with event-denoting complements if both aspectual and psychological verbs select for events. Crucially, given this difference in frequency distribution, if sentences with coercion verbs require the type-shifting operation to repair the type mismatch, then the nominal complements of psychological verbs should undergo coercion more frequently than the complements of aspectual verbs. This predicts more processing cost in psychological verbs than aspectual verbs—contrary to Katsika et al.'s findings. As it turns out, the set of "coercion verbs" investigated in previous studies encompasses a mixed bag of different verb classes that exhibit distinct linguistic distributions and processing profiles.

Furthermore, the SIH predicts that all sentences with aspectual verbs (AspVs) will be processed in a similar way. This includes sentences with an animate subject referent, such as (3.10a), and those with an inanimate subject referent, such as (3.11a). I will call the former "the AspVanimate configuration" and the latter "the AspVinanimate" configuration.

(3.10) a. Williams Shakespeare began the book.

b. ⇝ Agentive reading (along the eventive dimension): *Williams Shakespeare began reading/writing/etc. the book.*

c. ⇝ Constitutive reading (along the spatial/informational/etc. dimension): *The excerpt/words/etc. of Williams Shakespeare was the initial subpart of the book.*

(3.11) a. Shakespeare's excerpt began the book.

b. ⇝ Constitutive reading (along the spatial/informational/etc. dimension): *The excerpt/words/etc. of Williams Shakespeare was the initial subpart of the book.*

Without any biasing information in context, the AspVanimate configuration with an animate subject referent licenses both an agentive reading and a constitutive reading, while the AspVinanimate configuration with an inanimate subject referent typically engenders a constitutive reading only. However, both configurations involve dimension ambiguity; the constitutive reading in both configurations can be interpreted along multiple dimensions. That is, subject animacy alone is not the only factor that determines the interpretation for aspectual-verb sentences. The difference between the two lies in that the animate subject referent provides biasing information towards the agentive reading along the eventive dimension, in which the subject receives the agent role and the complement receives the theme role via the inverse thematic functions (see Section 3.2.3). The comparison between the two configurations

will be explicated in Chapter 5.

Second, hypothesizing that the cost associated with AspV sentences results from the resolution of dimension ambiguity, SIH predicts that biasing information in context (towards a specific dimension) will influence the interpretation of AspV sentences. For instance, following a context sentence (3.12a) that biases towards the agentive reading along the eventive or (3.12b) that biases towards the constitutive reading along the informational dimension, the AspV sentences (3.13) will receive an agentive reading when following (3.12a) and a constitutive reading when following (3.12b) respectively.

(3.12) a. [Context$_{:eventive}$]: Famous writers read extensively and habitually.

b. [Context$_{:informational}$]: Here are some anthologies that collect excerpts from well-known writers.

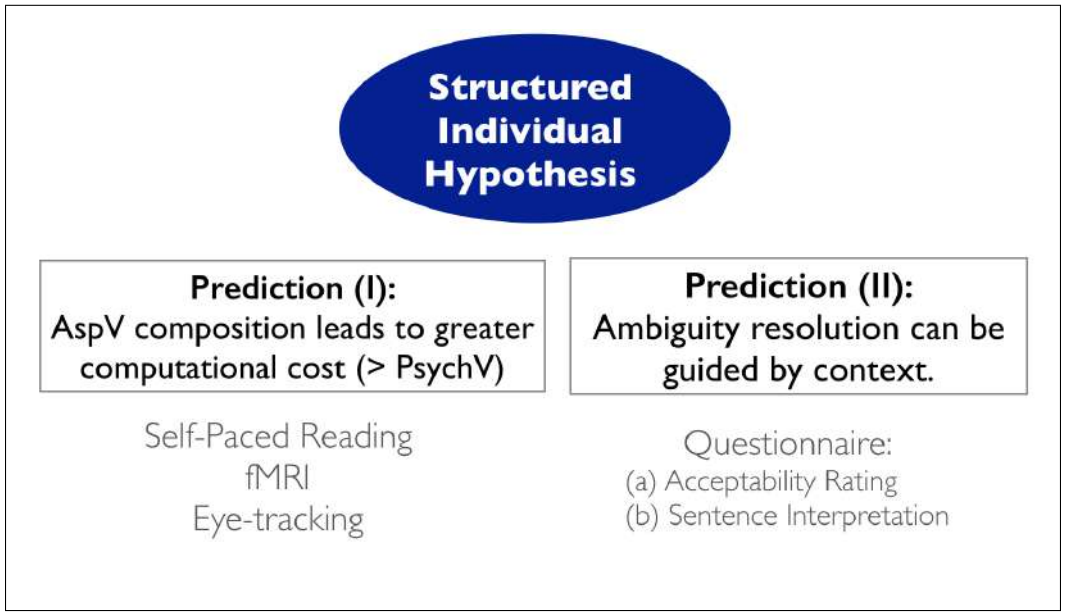
(3.13) Williams Shakespeare began the book. (= 3.10a)

Suppose that disambiguating contextual information guides the interpretation of the sentences, the SIH predicts that a disambiguating context will ease comprehenders' effort for ambiguity resolution and may increase acceptability of AspV sentences.

These predictions of the SIH have been tested respectively and will be presented in the following chapters. To forecast, Chapter 4 reports two experiments (self-paced reading & fMRI) that examine the first set of prediction: processing cost observed in previous studies of complement coercion is associ-

ated with aspectual verbs but not psychological verbs. Chapter 5 reports two experiments which show that all AspV sentences exhibit a similar processing profile regardless of subject animacy. Chapter 6 reports an eye-tracking experiment that examines whether context influences the processing of AspV sentences. Chapter 7 reports a questionnaire which shows that disambiguating contextual information guides the interpretation of AspV sentences and increases sentence acceptability. Figure 3.3 provides an overview of the SIH's predictions, along with the experiments reported in this dissertation that test the predictions. In the next chapter, I will present the first study that examine the first set of the SIH's prediction.

Figure 3.3: The predictions of the Structured Individual Hypothesis



Chapter 4

Complement coercion as the processing of aspectual verbs (Study 1)

4.1 Introduction

This study aims to examine the first prediction of the Structured Individual Hypothesis: processing sentences containing aspectual verbs will engender additional cost, because it involves (A) exhaustive activation of lexical dimension-functions encoded in the verbs (e.g. f_{space}, f_{time}, f_{info}, f_{event}), and (B) resolution of dimension ambiguity as the structured individual denoted by the complement can be mapped on to an axis along a variety of dimensions

(e.g. spatial, temporal, informational, eventive). To test this prediction, I conducted a self-paced reading experiment (Experiment 1) and a functional magnetic resonance imaging (fMRI) experiment (Experiment 2), along with a questionnaire (Pretest), reported in this chapter. On the Structured Individual Hypothesis, sentences with aspectual verbs should show prolonged reading times in Experiment 1 and distinct neural correlates in Experiment 2 associated with the two processes, as compared to psychological verbs.

4.2 Pretest: Rating questionnaire

A rating questionnaire was employed to ensure that the stimuli of the manipulated conditions are acceptable to native speakers.

4.2.1 Method

Participants: Forty native speakers of American English were recruited, all between the ages of 18-30 and without reading disabilities.

Materials: Fifty triplets of the three manipulated conditions were created, shown in Table 4.1.

Table 4.1: Study 1—Conditions and sample sentences

Condition	Example sentences
AspectualV	*Lady Gaga* ***started*** *this CD of American pop hits.*
EnjoyingV	..... ***enjoyed*** .....
LovingV	..... ***loved*** .....

Aspectual verbs (AspectualV) such as *start, begin, finish* were contrasted with psychological verbs such as *enjoy, endure, prefer* (EnjoyingV). These two verb classes are collapsed as "coercion verbs" in previous studies. I further introduced other psychological verbs such as *love* and *hate* as the control condition (LovingV).

On the Type-Shifting account, EnjoyingV and LovingV differ in that only the former exclusively selects for event-denoting complements whereas the latter does not (Pustejovsky, 1995, p.136)[1]. Therefore, on this account, only EnjoyingV but not LovingV will involve the type-shifting operation and induce processing cost comparable to AspectualV. On the Structured Individual hypothesis, the two sets of psychological verbs do not differ in selectional restriction and should pattern alike. The only difference is that the verbs in the EnjoyingV condition have been used as stimuli, treated as coercion verbs in previous studies of complement coercion, whereas those in the LovingV condition have not. The comparison is shown in Table 4.2

Table 4.2: Study 1—Predictions of the two hypotheses

Hypothesis	Prediction
Type-Shifting Hypothesis:	[AspectualV = EnjoyingV] > LovingV
Structured Individual Hypothesis:	[AspectualV] > [EnjoyingV = LovingV]

[1]Pustejovsky notes the subcategorization difference between the verbs *enjoy* and *like*, and indicates that "the semantics of these verbs are intact distinguished by the type of complement each selects." According to him, "the verb *enjoy* selects an event function, while *like* selects for an attitude towards any type." (1995: 135-6)

A hundred and fifty filler sentences were introduced. The whole set of the stimuli amounted to 300 sentences in total. These stimuli were divided into two scripts, and the participants were randomly assigned to one of them.

Procedures: For each sentence, the participants were asked to rate its acceptability on a 1∼5 scale (1=does NOT make sense; 5=makes sense) and answer a multiple-choice, multiple-answer question probing possible interpretations afterwards.

4.2.2 Results

The data of three participants were discarded because their responses were either undifferentiated or inconsistent. The rating scores showed that the critical sentences of the three conditions were within the acceptable range (> 3.5) in average (Table 4.3).

Table 4.3: Study 1—Results of rating questionnaire (Pretest)

Condition	Mean	Standard deviation
AspectualV	4.13	0.75
EnjoyingV	4.31	0.52
LovingV	4.80	0.22

A repeated measures Analysis of Variance (ANOVA) was performed to examine the differences among the conditions. Planned pairwise comparisons showed a significant difference between the AspectualV and LovingV conditions ($p<.001$) as well as between the LovingV and EnjoyingV conditions

(p<.001). Crucially, AspectualV and EnjoyingV conditions did not differ from each other (p<.082). In addition, I performed a reliability test, and results showed that the items within each condition exhibit high internal consistency (Cronbach's α: AspectualV=0.92; EnjoyingV=0.85; LovingV=0.72).

4.3 Experiment 1: Self-paced reading

Next, a self-paced reading experiment with the moving window paradigm was conducted. This aims to examine the behavioral patterns associated with the processing of aspectual verbs and psychological verbs.

4.3.1 Method

4.3.1.1 Participants

Twenty-eight native speakers of American English were recruited, all between the ages of 18-30 with normal vision. None of them had history of reading disabilities.

4.3.1.2 Materials

The materials were adapted from the pretest norming questionnaire. Each sentence was segmented into several windows as shown in the cells in Table 4.4. The windows of interest were the verb, complement head, and the two

subsequent segments (Head+1, Head+2).

Table 4.4: Study 1—Sample sentences and segmentation of Exp. 1

		Verb	Complement Head	Head+1	Head+2	
AspectualV	*Lady Gaga*	**started**	*this CD*	*of*	*American*	*pop hits.*
EnjoyingV	...	**enjoyed**	...	...	...	....
LovingV	...	**loved**	...	...	...	....

The verbs used and their frequencies in parentheses appeared in the stimuli are the following: AspectualVs included *start* (12), *begin* (13), *finish* (12), *continue* (9), *complete* (3), *end* (1). EnjoyingVs included *enjoy* (14), *prefer* (10), *favor* (9), *tolerate* (8), *endure* (8), *resist* (1). These verbs were taken from previous studies and have been used in Katsika et al. (2012). LovingVs included *love* (13), *like* (12), *dislike* (8), *hate* (5), *detest* (5), *approve of* (3), *be fond of* (2), *disapprove of* (1), *respect* (1).

The verbs in the three conditions were matched by reaction times from an independent lexical decision study carried out by DiNardo (2015). Results showed no difference in accessing times among AspectualVs (465.77ms), EnjoyingVs (454.46ms), and LovingVs (488ms), all $ps > .05$.

4.3.1.3 Procedures

Sentences were visually presented in black Courier New font in the center of a computer screen with a white background. The participants were instructed to read the sentences segment-by-segment at their own pace, in a fashion that allowed them to fully understand the meaning of the sentences. Every trial

began with a series of dash lines, with a "+" sign at the left edge of the sentence signaling the starting point. The participants began by pressing the spacebar, causing the first segment to appear. With the subsequent pressing, the next segment appeared, and the previous segment was replaced by a set of dashes. At the end of the sentence, the participants were presented a statement probing either the content or the acceptability of the sentence just read to ensure full comprehension. They responded by pressing the "Agree/Disagree" keys on the keyboard. A practice session was given beforehand; the participants had to reach 80% accuracy in the comprehension task before proceeding to the real trials.

4.3.2 Data analysis

The reading time (RT) data of all 28 participants were analyzed using the mixed model analysis, incorporating a fixed effect of condition (3 levels: AspectualV, EnjoyingV, LovingV) and random intercepts for subjects and items. None of the participants were excluded. Analyses were implemented in the R statistical environment with the *lme4* packages (Baayen et al., 2008; R Core Team, 2014). I evaluated the RT measurements by contrasting a model including condition as the fixed factor against a null model without it. For the pairwise comparisons, the p-values were corrected by the Tukey tests; the b values reported below represented the unstandardized coefficients. All

significant contrasts are reported.

4.3.3 Results

The accuracy of the comprehension task was 95.03%, indicating that the participants paid attention to the reading task and fully comprehended the sentences. Results of the reading times are reported in Table 4.5 and illustrated in Figure 4.1.

Table 4.5: Study 1—Exp. 1. Results of reading times (ms), standard errors in parentheses

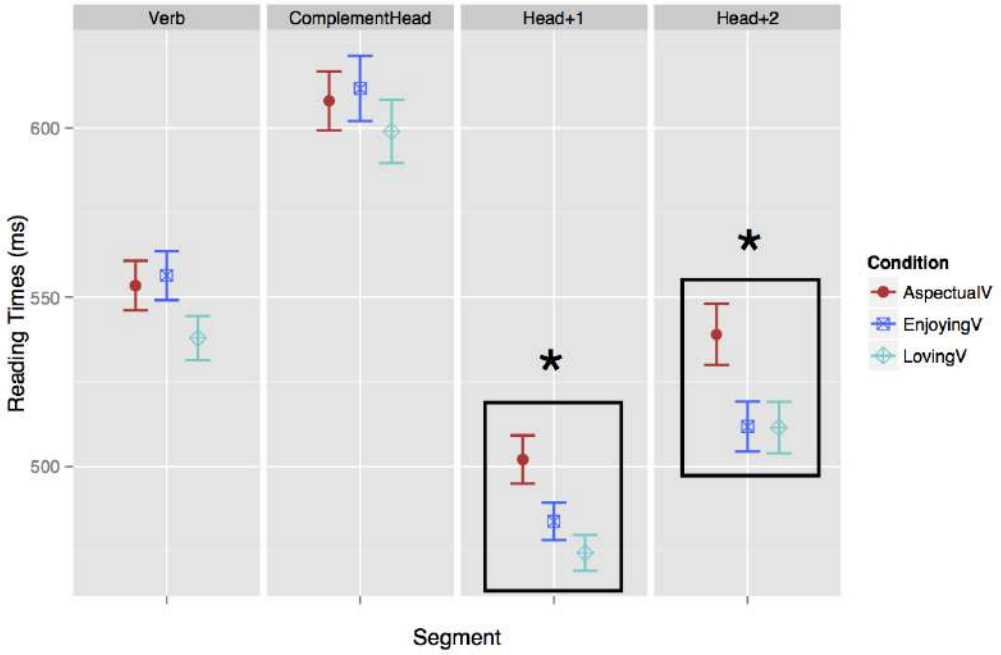
Condition	Verb	Complement Head	Head+1	Head+2
AspectualV	553.45 (7.30)	608.01 (8.70)	502.08 (7.13)	538.99 (9.07)
EnjoyingV	556.42 (7.26)	611.64 (9.61)	483.78 (5.55)	511.82 (7.38)
LovingV	537.93 (6.58)	598.98 (9.34)	474.50 (5.28)	511.49 (7.59)

The Verb region and the Complement Head region did not show a significant effect[2]. Significant differences among the conditions were found at the subsequent two segments, i.e. Head+1 and Head+2.

At the Head+1 region, a significant effect of condition was found ($\chi^2(2)=$ 14.315, $p < .001$). Pairwise comparisons (corrected) revealed that AspectualV induced significantly longer RTs than both EnjoyingV (b=18.301, $p<$ .036) and LovingV (b= 27. 581, $p<$ .001) respectively.

[2]The Verb region showed a marginal effect of condition ($\chi^2(2)=$ 5.475, $p <$ .0647), in the direction that that AspectualV/EnjoyingV > LovingV, but did not reach significance. This marginal effect went away at the subsequent Complement Head region ($\chi^2(2)$=1.445, p= .486)

Figure 4.1: Study 1—Exp. 1. Results of reading times (ms)



The Head+2 region exhibited the same pattern, showing a significant effect of condition ($\chi^2(2)= 11.197$, $p< .004$). Pairwise comparisons indicated that AspectualV induced significantly longer RTs than both EnjoyingV (b=27.172, $p< .011$) and LovingV (b=27. 504, $p< .010$) respectively.

Overall, results show that processing sentences with aspectual verbs engendered longer RTs than both sets of psychological verbs at the two segments following the complement head, while the two sets of psychological verbs did not differ from one another. These replicate Katsika et al.'s (2012) findings, and support Utt et al.'s dissociation of aspectual versus psychological verbs.

These results support the Structured Individual Hypothesis, which predicts that processing aspectual verbs, but not psychological verbs, engenders more cost because it involves resolution of dimension ambiguity. The pattern of longer RTs found after the complement head suggests that comprehenders make effort to determine the relevant dimension along which the sentence is to be interpreted among multiple ones (e.g. spatial, eventive, informational).

4.4 Experiment 2: fMRI

The neural correlates associated with complement coercion have been found in three distinct brain regions, mentioned in Chapter 2 (Section 2.4.1.2). Piñango and Zurif (2001) found that Wernicke's aphasics have difficulty comprehending sentences with complement coercion. Pylkkänen and McElree

(2007), using MEG, found that coercion sentences induced activity in anterior inferior midline filed (AMF), which might be generated in ventral medial prefrontal cortex (vmPFC). More recently, Husband et al.'s (2011) fMRI study reports that processing complement coercion recruited Brodmann area (BA) 45 in left inferior frontal gyrus (LIFG). Despite of the discrepant findings, all these studies attribute the effect to the implementation of the type-shifting operation to repair the mismatch between the event-selecting verb and the entity-denoting complement, so as to yield an eventive interpretation.

It should be pointed out that the studies mentioned above are subject to the problem of heterogeneous stimuli. The set of coercion verbs contain a mixed bag of aspectual verbs, psychological verbs, and others (e.g. *master, try*). This is problematic as suggested by Katsika et al. (2012), Utt et al. (2013), and Experiment 1 reported in Section 4.3.

The present experiment differs from previous studies in isolating aspectual verbs as a unified classes distinct from psychological verbs, capturing the complement coercion effect as the processing of aspectual verbs. In Experiment 2, I aim to investigate the neural basis corresponding to the two processes proposed by the Structured Individual Hypothesis—(A) exhaustive retrieval of lexical dimension functions in the verb, and (B) resolution of dimension ambiguity.

4.4.1 Method

4.4.1.1 Participants

Sixteen native speakers of American English between the ages of 18-30 were recruited. All were right-handed and had no reading disabilities or history of neurological disorders.

4.4.1.2 Materials

The stimuli were adopted from Experiment 1, with a different set of fillers. The whole set of stimuli contained 50 triplets of the 3 conditions and 150 fillers, amounting to 300 sentences in total.

4.4.1.3 Paradigm

An event-related paradigm was employed. Sentences were visually presented segment-by-segment as in Experiment 1, each segment lasting for 500ms. The participants were queried with a yes/no comprehension question at the end of the sentence just read for 75% of the trials. The question queried the content of various parts of the sentence to avoid strategic reading. There was a 500 ms interval between the sentence-final word and the comprehension question following the sentence.

The total 300 sentences were divided into 10 runs. Sentences within each

run were pseudo-randomized such that no successive sentences were of the same condition, and the order of the runs were randomized for each participant. Each run contained 30 sentences, lasting for 5 minutes 33 seconds with the inclusion of the machine connection delay.

4.4.1.4 Imaging acquisition

Anatomical Measurements: The fMRI experiment was performed on a Siemens Sonata; 3T whole body MRI scanner (Erlangen, Germany). Each session began with a 3 plane localizer (20 seconds) followed by a Sagittal localizer, inversion recovery T1 weighted scan (TE/TR=2.61/285, matrix 192x192) 2.5mm thick, FOV=220 mm, bandwidth = 501 Hz/pix, 51 slices. This acquisition was used to define the ACPC line for prescription of the anatomic T1 images and functional images in the following series. Axial oblique T1 weighted images were recorded next, to be used as basic anatomic images that were later registered to the subjects own 3D high resolution acquisition obtained at the end of the study, which in turn was then registered to a high resolution reference brain for multi-subject analysis.

Functional Measurements: During the task, the event-related functional MRI was conducted, using gradient echo EPI blood oxygenation level dependent (BOLD) contrast. The scanner was set to trigger the stimulus presentation program, which enabled the image acquisition to be synchronized with

the stimulus presentation. The EPI imaging parameters were TE =30 ms, TR=956 ms, matrix 84 x 84, 51 slices. A high resolution volume acquisition then followed: 3D MPRAGE (the fMRI data within subjects was registered to this brain volume and this brain volume was then registered across subjects into a common 3D brain space using in house developed registration software (Studholme et al., 2001). The parameter was 256x256 acquisition matrix, FOV=256mm, bandwidth = 179 Hz/pix, TE=2.77 ms, TR=2530 ms, 2 averages, 176 slices.

4.4.2 fMRI data analysis (event-related)

All data were converted from Digital Imaging and Communication in Medicine (DICOM) format to analyze format using XMedCon (Nolfe et al., 2003). During the conversion process, the first 6 images at the beginning of each of the 10 functional runs were discarded to enable the signal to achieve steady-state equilibrium between radio frequency pulsing and relaxation, leaving 315 images per slice per run for analysis. Functional images were motion-corrected with the Statistical Parametric Mapping (SPM) 5 algorithm (www.fil.ion.ucl.ac.uk/spm/software/spm5) for three translational directions (x, y, z) and three possible rotations (pitch, yaw, roll). Trials with linear motion that had a displacement exceeding 1.5 mm or rotation exceeding 2 degrees were rejected. The data from one participant were excluded because of severe head movements. All further analyses

were performed using BioImage Suite (Papademetris et al., 2006).

Event segmentation: Individual subject data were analyzed using a General Linear Model (GLM) on each voxel in the entire brain volume with regressors specific for each task. For each of the three conditions (AspectualV, EnjoyingV, LovingV), there were two regressors for the two events corresponding to the two hypothesized processes for comprehending sentences with aspectual verbs. **Event 1** included the onset of the subject noun phrase until the offset of the main verb. On the Structured Individual Hypothesis, comprehenders should exhaustively retrieve the dimension functions lexically encoded in the verbs. **Event 2** included the onset of the complement noun phrase until the offset of the sentence-final word. During this window, comprehenders are hypothesized to resolve dimension ambiguity in order to determine the relevant dimension along which the complement denotation is construed as an axis.

Table 4.6: Study 1—Exp. 2. Event segmentation in fMRI data analysis (cells representing the stimuli presentation)

Condition	**Event 1** (Subject + Verb)		**Event 2** (Complement ∼ Sentence-final)			
AspectualV	*Lady Gaga*	*started*	*this CD*	*of*	*American*	*pop hits.*
EnjoyingV	*Van Gogh*	*enjoyed*	...	...	...	...
LovingV	*Van Gogh*	*loved*	...	...	...	...

The resulting beta images for each task were spatially smoothed with a 6 mm Gaussian kernel to account for variations in the location of activation across subjects. The output maps were normalized beta-maps, which were in the acquired space (2.5 x 2.5 x 2.5mm).

To take these data into a common reference space, three registrations were calculated within the Yale BioImage Suite software package. The first registration performed a linear registration between the individual subject raw functional image and that subject's 2D anatomical image. The 2D anatomical image was then linearly registered to the individual's 3D anatomical image. The 3D differs from the 2D in that it has a 1x1x1 mm resolution whereas the 2D z-dimension is set by slice-thickness and its x-y dimensions are set by voxel size. Finally, a non-linear registration was computed between the individual 3D anatomical image and a reference 3D image. The reference brain used was the Colin27 Brain (Holmes et al., 1998) in Montreal Neurological Institute (MNI) space (Evans et al., 1993). All three registrations were applied sequentially to the individual normalized beta-maps to bring all data into the common reference space.

Data were corrected for multiple comparisons by spatial extent of contiguous suprathresholded individual voxels at an experiment-wise $p< .05$. In a Monte Carlo simulation within the AFNI software package and using a smoothing kernel of 6mm and a connection radius of 4.33mm on 2.5 x 2.5 x 2.5mm voxels, it was determined that an activation volume of 183 original voxels (4953 microliters) satisfied the $p< .05$ threshold. Clusters were created for each of the subtractions. Each cluster was identified with a region label, and then associated with additional numeral labels corresponding to Brodmann areas.

4.4.3 Results

4.4.3.1 Behavioral results

The overall accuracy of the comprehension task was 88.6%, suggesting that the participants were paying attention to the task. The response times for the questions are shown in Table 4.7 and Figure 4.2. A mixed model analysis was performed, incorporating condition as the fixed factor, subjects and items as the random factors. Results show no significant effect of condition in response times. ($\chi^2(2) = 1.238$, $p< 0.54$).

Table 4.7: Study 1—Exp. 2. Response times (ms) of the comprehension questions

Condition	Mean (ms)	Standard Error	SD
AspectualV	1684.18	25.06	590.9
EnjoyingV	1722.11	25.08	582.23
LovingV	1651.49	24.09	547.17

4.4.3.2 Imaging results

At Event 1 (Subject + Verb), AspectualV engendered more activity in Wernicke's area (BA40), bilateral BA 6, 24, 7, and primary sensory areas, as compared to EnjoyingV. At Event 2 (Complement∼Sentence-final word), AspectualV engendered more activity in left inferior frontal cortex (LIFC), including BA44, 44, 47, left insula, bilateral BA6, right inferior frontal gyrus, right BA8, and primary visual cortex, as compared to the control LovingV. These are shown in Figure 4.3.

Figure 4.2: Study 1—Exp. 2. Response times (±1 SE) of the comprehension questions

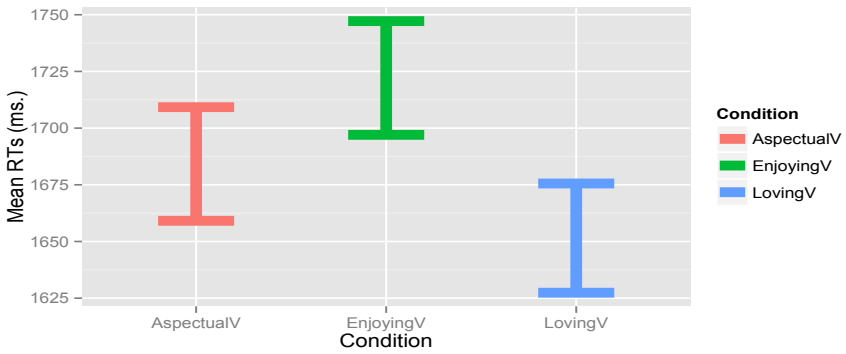


Figure 4.3: Study 1—Results of Exp. 2. *Left*: AspectualV > EnjoyingV at Event 1 (Subject + Verb). *Right*: AspectualV > LovingV at Event 2 (Complement∼S-final).

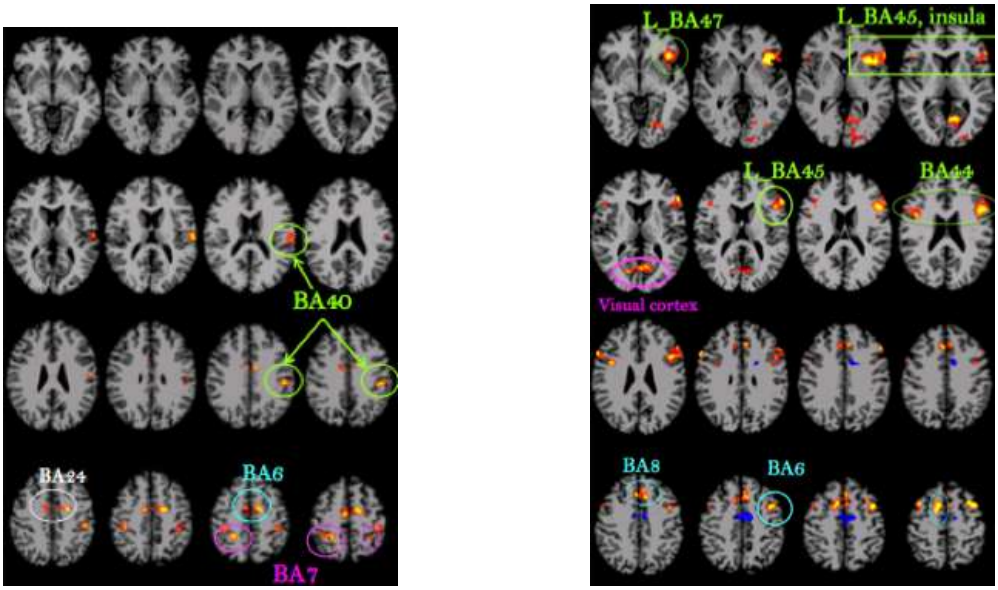


Table 4.8: Study 1—Results of Exp. 2. The differential activated regions of AspectualV > EnjoyingV at Event 1.
(L=left, R=right, AntCingulate= anterior cingulate, PrimSensory=primary sensory cortex, PrimMotor=primary motor cortex.)

Region	Volume (mm^3)	Mean T-Value	Max T-Value	MNICoord Max
L_BA40	*2168*	*2.52033*	*3.55617*	*-48,-27,39*
L_BA7	*2761*	*2.6424*	*4.39206*	*-30,-54,69*
Medial_BA6_AntCingulate	7735	2.53566	4.9745	-18,-6,45
R_BA7_PrimSensory	7377	2.61813	5.71615	27,-45,51
L_PrimSensory_PrimMotor	*4081*	*2.44217*	*4.02101*	*-60,-15,12*

Table 4.9: Study 1—Results of Exp. 2. The differential activated regions of AspectualV > LovingV at Event 2.
(L=left, R=right, IFG=inferior frontal gyrus, AntCingulate=anterior cingulate)

Region	Volume (mm^3)	Mean T-Value	Max T-Value	MNICoord Max
L_Insula	*1473*	*3.28132*	*6.8164*	*-36,18,0*
L_IFG_BA47_BA45_BA44	*11565*	*2.75671*	*6.73777*	*-37,18,0*
L_lateral_BA6	*3718*	*2.83759*	*5.80195*	*-39,9,51*
R_AntCingulate	*248*	*2.63028*	*3.68964*	*-6,30,30*
R_IFG_BA44_BA45	*3135*	*2.57169*	*4.19239*	*60,18,27*
R_medial_BA6_BA8	*5647*	*2.70576*	*5.40665*	*15,0,54*
R_lateral_BA6_BA8	*1080*	*2.53198*	*5.34182*	*39,12,54*
Medial_BA6_AntCingulate	4958	-2.59486	-4.74066	-18,-15,78
PrimaryVisual_BA18_BA19	6148	2.54813	4.2163	-12,-63,6

4.5 Discussion

Summarizing the results: In Experiment 1 (Self-paced reading), results show that sentences with aspectual verbs engendered longer RTs than the two sets of psychological verbs (i.e. EnjoyingV and LovingV) after the complement head was encountered, while the two sets of psychological verbs did not differ from each other. In Experiment 2 (fMRI), results show that aspectual verbs preferentially recruited BA 40 (part of Wernicke's area) during the composition of the subject and the verb, as well as LIFC after the complement was encountered.

The findings of Exp.1 replicate Katsika et al. (2012), whose eye-movement results show additional cost for aspectual verbs as compared to psychological verbs. Results also support Utt et al.'s (2013) suggestion that aspectual verbs and psychological verbs should be dissociated based on their distribution in corpus (Chapter 2, Section 2.4.2). Aspectual and psychological verbs differ not only in linguistic behaviors but also in processing profiles, suggesting that the set of coercion verbs identified in previous work are semantically heterogeneous. The fact that the two verb types show different processing profiles weakens the Type-Shifting Hypothesis. If event-selecting verbs, including both verb types, require type-shifting the entity-denoting complement into an event to resolve the mismatch, then both should pattern alike in engendering

additional cost—contrary to the findings.

On the other hand, the Structured Individual Hypothesis predicts that the processing cost associated with aspectual verbs results from (A) exhaustively retrieving the dimension functions encoded in the verb, and (B) resolution of dimension ambiguity in context to interpret the complement and consequently, the sentence. Upon encountering the complement, comprehenders conceive it as a structured individual. While the structured individual denoted by the complement can be mapped along many dimensions, comprehenders attempt to identify a context-relevant dimension to interpret the sentence. I take the longer RTs found after the complement head in Exp. 1 as reflecting the process of determining a relevant dimension in context (along which the structured individual denoted by the complement is mapped onto) among multiple alternatives.

This explanation is supported by the findings of Traxler et al. (2005), which seem puzzling for the Type-Shifting Hypothesis. Their eye-tracking experiment shows that the cost of complement coercion was attenuated when following a prior context sentence that either contained the same coercion verb or explicitly mentioned the event structure. When either "*started a book*" or "*read a book*" is given in the context sentence (4.1a), there was no difference between "*started the book*" and "*read the book*" in the subsequent target sentence (4.1b).

(4.1) a. [Context]: *The student started/read a book....*

b. [Target]: *Before he started/read the book....*

—Traxler et al. (2005)

The Structured Individual Hypothesis captures this finding naturally: The context sentence provides information that determines or biases toward a specific dimension (likely the eventive one in this example), and therefore lessens the effort of resolving dimension ambiguity in the target sentence. Consequently, the processing cost was attenuated.

Let me now turn to Experiment 2 (fMRI). The imaging results revealed that aspectual verbs preferentially recruited left posterior superior temporoparietal cortex (part of Wernicke's area) when the verb was encountered and LIFC (Broca's area) after the complement was encountered, as compared to the psychological verb conditions. On the Structured Individual Hypothesis, the posterior temporal activity likely reflects the exhaustive retrieval of lexical functions stored in aspectual verbs. Previous studies have suggested that this area is engaged in the retrieval of lexico-semantic representations (Badre et al., 2005; Binder et al., 2009; Damasio et al., 1996; Hickok and Poeppel, 2004, 2007; Humphries et al., 2007; Lau et al., 2008). Aspectual verbs encode multiple lexical-dimension functions (e.g. f_{space}, f_{info}, f_{time}, f_{event}), and could thus involve more effort for retrieving lexical information, as compared to psychological verbs. The activation of discrete lexical functions might be a

measure of complexity that correlates with the strength of activity in the posterior temporal region. This interpretation is consistent with the studies of Shapiro et al. (1993)[3] and Shetreet et al. (2016)[4], which show that Wernicke's area is specifically involved in exhaustive retrieval of lexical meanings during comprehension.

Noticeably, sentences with aspectual verbs recruited the precuneus (BA7, marked by the circle on the left panel of Figure 4.3) at Event 1 (Subject + Verb), a finding of interest to the SIH. This region has been suggested to support spatio-temporal tasks, such as processing visual-spatial information in perception and memory, especially the spatial representation of sequential movements (Cavanna and Trimble, 2006; Fletcher et al., 1995; Wallentin et al., 2008). The activation of the precuneus, though unexpected initially, is consistent with the SIH: It is likely to reflect the axial conceptualization associated with the dimension functions, one of which will map the structured individual

[3]Shapiro et al. (1993) compared transitive verbs (e.g., *fix*), which have only a two-place argument structure (Agent, Theme) vs. dative verbs (e.g., *send*), which allow both a two-place and three-place (Agent, Theme, Goal) argument structure. They also compared two-complement verbs (e.g., *expect*), which allow both an (Agent, Theme) and an (Agent, Proposition) frame vs. four-complement verbs (e.g., *discover*), which allow (Agent, Theme), (Agent, Proposition), (Agent, Exclamation), and (Agent, Interrogative). Results from their cross-model lexical decision task revealed that normal individuals and Broca's aphasics showed shorter reaction times for transitive verbs than dative verbs, and faster for two-complement verbs than four-complements verbs. Yet such verb type effect was not observed in Wernicke's aphasics, suggesting that Wernicke's area supports the retrieval of verbs' thematic representations.

[4]Shetreet et al. (2016) found that multiple-option verbs—referring to the subcategorization frames and semantic frames—induced stronger activation in left posterior superior temporal gyrus (STG) than one-option verbs. For instance, the multiple-option verb *remember* can take a theme/NP (e.g., *Dan remembered Jenny.*) or a proposition/embedded clause (e.g., *Dan remembered that Jenny is coming tonight.*), whereas the one-option verb *punish* takes only a theme/NP (e.g., *Dan punished Jenny.*).

onto an axis. The notion of structured individual is connectable to spatial conceptualization, and presumably provides a window that allows some linkage between spatial representation and lexical conceptual structure, as suggested by the Parallel Architecture (Jackendoff, 2007, 2010). The recruitment of the precuneus and the linkage to spatial conceptualization cannot be naturally connected to the Type-Shifting Hypothesis however.

The activity of LIFC and left insula observed at Event 2 (Complement∼Sentence-final) for aspectual verbs is likely to reflect the processing of resolving the dimension ambiguity in order to construe the structured individual denoted by the complement. This region has been suggested to support ambiguity resolution (Badre et al., 2005; Krain et al., 2006; Lau et al., 2008; Rodd et al., 2010). The insula has been reported to engage in the processing of time. Magnani et al.'s (2014) fMRI study employed prismatic adaptation (PA); the participants were asked to indicate the time intervals as perceived in a time reproduction task. The PA-induced rightward aftereffect resulted in an overestimation of time intervals whereas the leftward aftereffect resulted in an underestimation of time intervals. This task induced increased activity in left anterior insula and left superior frontal gyrus *after* PA, as compared to activity before PA. Magnani et al. therefore suggest that these areas underlie the spatial manipulation of the representation of time. In the case of aspectual verbs, it is possible that the activity in left insula reflects the schematic representation of structural configuration, such as the precedence relation on the

axis, which could be temporal-spatial.

Another activated region worth noting is the supplementary motor area (BA6), recruited for aspectual verbs at Event 2. Previous studies have suggested that this area is involved in action planning or event sequencing (Crozier et al., 1999), action simulation, the generation of ordinally structured sequences (Stadler et al., 2011), and the updating of spatial information (Tanaka et al., 2005). Based on these findings, the activation of BA 6 is not surprising to the SIH. It could be associated with the sequential action planning along the eventive dimension—one that yields the agentive reading for aspectual-verb sentences, such as "*John began (reading) the book (and then he jotted down some notes...)*."

Overall, the imaging results replicate and extend previous studies regarding the neural correlates associated with complement coercion. Let aside the heterogeneity of the verb types used in previous studies, Experiment 2, by using an event-related paradigm, shows that the two regions found in Piñango and Zurif (2001) and Husband et al. (2011) are both involved, yet at different time windows. Wernicke's area is recruited during the composition of the subject and the aspectual verb, whereas LIFC is recruited when the complement denotation is accessed. These neural correlates and their timing profiles are consistent with the Structured Individual hypothesis. To the Type-Shifting Hypothesis, this distributed recruitment pattern comes as a surprise, as it would expect one cortical region that supports the implementation of the

type-shifting operation, contrary to the findings.

4.6 Conclusion

In this chapter, I showed that (1) aspectual verbs form a unified verb class and should be dissociated from psychological verbs based on both behavioral and neurological measurements in comprehension, and therefore (2) the complement coercion effect should be viewed as the processing of aspectual verbs.

The results of the two experiments evidence that processing aspectual verbs engenders additional cost and neural activity during real-time comprehension, as compared to psychological verbs. Aside from the the distinct linguistic distributions of two verb types (Piñango and Deo, 2015; Utt et al., 2013), their behavioral and neurological profiles differ as well. These findings validate the first prediction of the SIH indicated in Chapter 3 (Section 3.3.2): When the two verb types are dissociated, only aspectual verbs are costly. Therefore, the complement coercion effect is better understood as processing of aspectual verbs, since not all coercion verbs engender cost.

One might want to retain the Type-Shifting Hypothesis by claiming that the type-shifting operation applies only to aspectual verbs and hence only these verbs engender processing cost. However, cases like "*The short postscript ends the novel.*" undermine this solution, because aspectual verbs do not necessarily select for eventive complement, contrary to the assumption of the

Type-Shifting Hypothesis. In such sentences with aspectual verbs, there is no type mismatch and therefore the type-shifting operation should not be needed.

The pattern suggests a more straightforward incremental semantic composition that is rooted in lexical representation and constrained by contextual information. When a lexical item (e.g. an aspectual verb) is retrieved from long-term memory, its encoded lexical information, such as lexical functions and selectional restrictions, come with it automatically—the process of lexical retrieval as suggested by the SIH. With this lexical representation carried in short-term memory, the processor seeks to fulfill the requirements imposed by each lexical item and predicts what will come next (e.g., the conceptualization of a structured individual). If the composition of lexical representations gives rise to multiple plausible interpretations, the processor will determine an appropriate one supported by the context. This corresponds to the process of ambiguity resolution as suggested by the SIH.

As compared to the Type-Shifting Hypothesis, the view suggested by the SIH is more conservative in that no semantic operation extraneous to lexical representation is postulated to account for the complement coercion effect. What take place during real-time comprehension (as in the processing aspectual verbs) are the fulfillment of the requirements imposed by lexical semantics, and the contextual search for obtaining an appropriate interpretation for the sentence.

Study 1 reported in this chapter verifies the first prediction of the SIH

that AspV sentences engender more processing cost than PsychV sentences. However, sentences tested in this study all contain animate subject referents. Will AspV sentences with an inanimate subject referent (e.g., *The prologue begins the novel.*) pattern like those with an animate subject referent (e.g, *The writer began the novel.*)? The SIH predicts that all sentences with aspectual verbs will exhibit a similar processing profile because they all involve lexical retrieval of the verb's functions and dimension ambiguity (recall that a constitutive reading can be interpreted along multiple dimensions). In the next chapter, I report Study 2 that tests this prediction.

Chapter 5

Processing of aspectual verbs in composition with animate and inanimate subjects (Study 2)

5.1 Introduction

Study 1 examines sentences with aspectual verbs composed with animate subject referents. However, aspectual verbs (AspVs) can combine with an inanimate subject referent as well; not only *"The writer began the book"*, but also *"The prologue began the book"*. The question asked in this chapter is whether subject animacy factors in the processing of AspV sentences. The Structured Individual Hypothesis (SIH) predicts a similar pattern for AspV

sentences with animate and inanimate subject references, because all AspV sentences involve dimension ambiguity without context. In this chapter, I report two experiments that test this prediction of the SIH.

5.1.1 Two configurations in which aspectual verbs appear

Consider sentence (5.1a), in which an aspectual verb is composed with a subject that denotes an animate entity. I will call this combination "the AspVanimate configuration" henceforth. This configuration often gives rise to an agentive interpretation, as in (5.1b), in which the animate subject referent is conceived as an agent and the complement is conceptualized as a structured individual construed as an axis along the eventive dimension.

(5.1) a. Van Gogh began the collection of oil paintings.

b. ⇝ Van Gogh began *viewing/editing/etc...* the collection of oil paintings.

c. ⇝ *Van Gogh's painting* was the initial subpart of the collection of oil paintings.

Sentence (5.1a) can also give rise to a constitutive reading, such as (5.1c), in which the subject denotation is conceived as a subpart of the complement denotation. In this reading, the complement is construed as a structured

individual, which can be mapped along multiple dimensions. For example, "the collection" can be conceived as a physical entity, consisting of X pages, along the *spatial* dimension. The sentence means that Van Gogh's painting is printed on the first page of the physical entity denoted by the complement. On the other hand, the complement can be construed as a structured individual mapped along the *informational* dimension. In this case, the complement is conceived as a structured body of informational content, consisting of X paintings (a painting is a piece of informational content). The sentence means that Van Gogh's painting is the first painting of this collection of oil paintings.

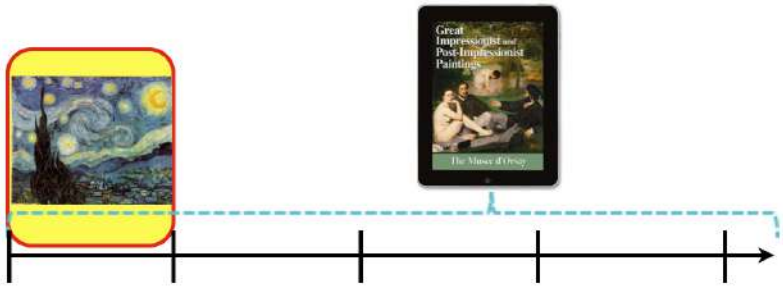
Therefore, the composition of an aspectual verb and an animate subject referent gives rise to a semantic ambiguity between an agentive reading and a constitutive reading. Comprehenders have to determine in context an appropriate reading along a contextually relevant dimension.

Let's now look at cases where an aspectual verb is composed with an animate subject referent, such as (5.2a). I will call it "the AspVinanimate configuration" henceforth. Such configuration gives rise to a constitutive reading, in which the subject denotation is mapped onto a specific subpart of the axis construed by the complement, as in (5.2b), illustrated in Figure 5.1.

(5.2) a. "Starry Night" began the collection of oil paintings.

b. $\rightsquigarrow$ "Starry Night" was the initial subpart of the collection of oil paintings.

Figure 5.1: The constitutive reading of an AspV sentence



Without any contextual information, the constitutive reading is underspecified with respect to the exact dimension along which the complement denotation is construed. And just like the constitutive reading of the AspVanimate configuration, the constitutive reading like (5.2b) for the AspVinanimate configuration can be mapped onto different dimensions (e.g. spatial or informational), yielding distinct interpretations.

The point that I want to emphasize here is that both the AspVanimate and AspVinanimate configurations exhibit dimension ambiguity. On the Structured Individual Hypothesis, processing all sentences with aspectual verbs will involve resolution of dimension ambiguity, regardless of subject animacy. Hence, both configurations will require additional processing effort, because the structured individual denoted by the complement can be construed along multiple dimensions in both cases.

5.1.2 The superset relation between the two configurations of aspectual verbs

The above section demonstrates that the two configurations of AspV sentences allow a constitutive reading along multiple dimensions. Nevertheless, in the AspVanimate configuration, the animate subject referent (preferably) yields an agentive reading along the eventive dimension *in addition to* the constitutive reading along multiple dimensions. This means that the set of

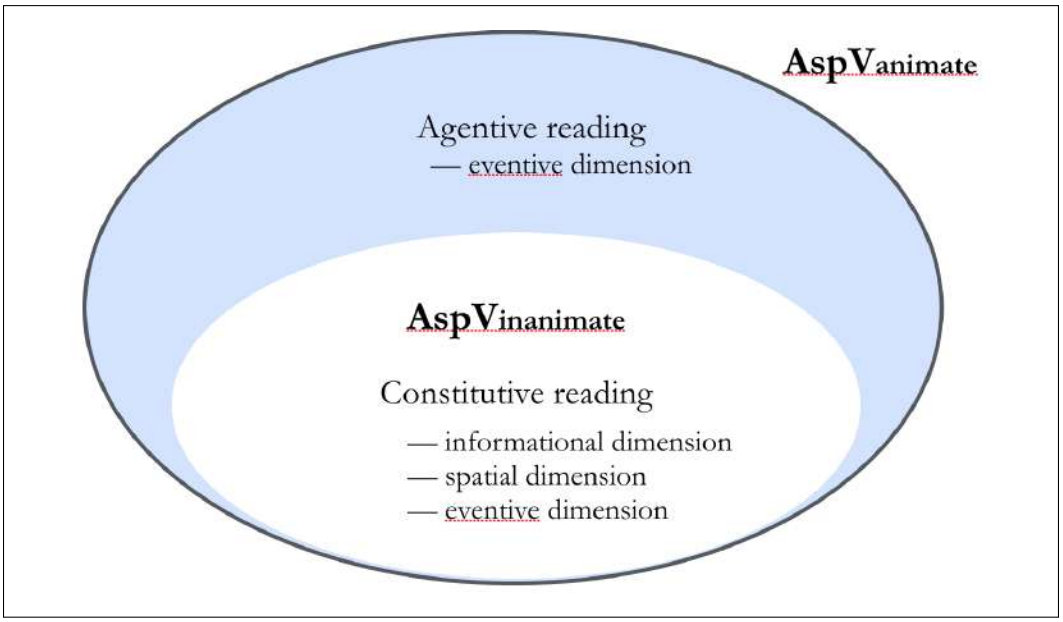
dimension readings available in the AspVanimate configuration constitute a *superset* of the readings available in the AspVinanimate configuration, i.e., AspVanimate $\supseteq$ AspVinanimate, as shown in Figure 5.2.

On the Structured Individual analysis (Piñango and Deo, 2015), the agentive reading along the eventive dimension, such as "Van Gogh began *browsing* the collection," involves the inverse thematic functions, f_{ag_i} and f_{th_i} that map the subject and complement denotations to an event that they are participants of at a reference interval i. Such mapping is absent in the constitutive reading, as in "Van Gogh's painting began the collection." It is of interest here to see if this difference between the two configurations can be observed in real-time comprehension. While both AspVanimate and AspVinanimate sentences involve dimension ambiguity, the former might engender a processing pattern distinguishable from the latter.

5.1.3 Components influencing the interpretations of AspV sentences

The interpretation of an AspV sentence is influenced by two components: (a) the control relation of the participants (individuals denoted by the subject and the complement), and (b) the choice of dimension. According to Piñango (to appear), *control relation* refers to "the degree of asymmetric power between the potential participants in a relation." Control asymmetry refers to a situa-

Figure 5.2: AspVanimate ⊇ AspVinanimate in terms of plausible dimension readings



tion in which a participant can exercise power over another participant. In the case of AspV sentences, an animate subject referent (as compared to an inanimate referent) is more likely to be perceived as a controller (i.e., agent), and hence could bias towards the agentive reading along the eventive dimension.

Another component that determines the interpretation for an AspV sentence is the choice of dimension. As indicated in Section 5.1.1, a constitutive reading can be interpreted along multiple dimensions. For either configuration of AspV sentences, the exact dimension reading depends on which dimension is chosen in context to interpret the sentence. The ultimate reading is a combination of the two components, (a) and (b), shown in Table 5.1.

Table 5.1: Two factors affecting the interpretations of AspV sentences

(a) Control relation between the participants	(b) Choice of Dimension
Asymmetric → Agentive reading	Eventive Spatial
Symmetric → Constitutive reading	Informational Temporal

Notice that although an animate subject referent can bias towards the agentive reading, ApsVanimate sentences do not necessarily yield such reading. In (5.3), the subject typically denotes an animate entity, but the sentence gives rise to a constitutive reading along the spatial dimension.

(5.3) The little girl began the line.

On the other hand, the eventive dimension by itself does not necessarily yield an agentive reading neither. This is demonstrated in (5.4), which gives rise to

a constitutive reading along the eventive dimension.

(5.4) The prayer finished the banquet.

More examples that illustrate the interaction of subject animacy and dimension are provided in Table 5.2.

Table 5.2: The interaction between subject animacy and dimension

Composition with Subject	**Dimension**	Reading	Example
Animate Subj + AspV	eventive	agentive	*Van Gogh$_{(agent)}$ began the collection.*
	informational	constitutive	*Van Gogh$_{(content)}$ begins the collection.*
	spatial		*The little girl began the queue.*
Inanimate Subj + AspV	eventive	agentive	*The printer started my paper.*
		constitutive	*The prayer finished the banquet.*
	informational		*A short postscript ends the novel.*
	spatial		*The famous perch begins the trail.*
	temporal		*The December solstice begins the winter.*

These cases suggest that neither subject animacy or dimension alone can predict the kind of reading (agentive vs. constitutive). Rather, the specific interpretation of sentences with aspectual verbs is determined by the control relation between the participants—which could be biased, but not determined by subject animacy—and the choice of dimension in context *in combination.*

Study 2 extends Study 1 to investigate the processing of the two configurations of aspectual verbs (AspVanimate, AspVinanimate). To this end, I introduce a condition in which an aspectual verb is composed with an inanimate subject referent. The Structured Individual Hypothesis predicts that all sentences with aspectual verbs, regardless of subject animacy, will engender additional processing cost because they all involve dimension ambiguity.

I test this prediction via a self-paced reading experiment (Experiment 1) and an fMRI experiment (Experiment 2), presented below.

5.2 Experiment 1: Self-paced reading

This experiment examines the behavioral patterns of the two configurations of aspectual verbs. It is predicted that they will be processed in a similar fashion in real-time comprehension, because both involve dimension composition of the verb and the complement as well as semantic ambiguity.

5.2.1 Method

5.2.1.1 Materials

Four conditions were employed, all contained entity-denoting complements (Figure 5.3). The two conditions with aspectual verbs differed minimally in subject animacy. The AspVanimate condition contained a subject that typically refers to an animate entity (e.g. *Van Gogh*). The AspVinanimate condition contained a subject that typically refers to an inanimate entity (e.g. *"Starry Night"*). These two conditions of aspectual verbs were contrasted with two sets of psychological verbs as controls, adopted from Study 1 (referred to as EnjoyingV and LovingV). Since it has been established that the two sets of psychological verbs do not differ from each other in real-time pro-

cessing in Study 1, I will call them "Control$_{PsychV1}$" and "Control$_{PsychV2}$" in this study. The former contained psychological verbs that are identified as coercion verbs by the Type-Shifting Hypothesis, such as *enjoy, endure, prefer*; the latter contained other psychological verbs such as *love, hate, like*. The SIH does not differentiate the two sets of psychological verbs; both of them serve as controls.

Table 5.3: Study 2—Conditions and sample sentences

Condition		Verb	Comp.Head	Comp.+1	Comp.+2	
AspVinanimate	*"**Starry Night**"*	***starts***	*the collection*	*of*	*impressionist*	*oil paintings.*
AspVanimate	***Van Gogh***	***started***	...	...	...	...
Control$_{PsychV1}$	*Van Gogh*	***enjoyed***	...	...	...	...
Control$_{PsychV2}$	*Van Gogh*	***loved***	...	...	...	...

Fifty quadruples of the four conditions were created and one hundred fillers were introduced, yielding a total of 300 sentences. The sentences of the AspVanimate, Control$_{PsychV1}$, and Control$_{PsychV2}$ conditions were the same as in Study 1. The sentences of the AspVinanimate condition were normed along with the other three conditions in a pretest rating questionnaire with a 1∼5 scale to ensure that the stimuli were acceptable to native speakers. Results show that all sentences were within the acceptable range (> 3.7) in average to native speakers of American English (n=37).

Table 5.4: Study 2—Results of the rating questionnaire. (5=makes sense; 1=does NOT make sense.) (n=37)

Condition	Mean	SD	se
AspVanimate	4.13	0.75	0.04
AspVinanimate	3.73	0.84	0.05
Control$_{PsychV1}$	4.31	0.52	0.04
Control$_{PsychV2}$	4.80	0.23	0.02

5.2.1.2 Participants

Twenty-eight native speakers of American English were recruited[1]. All were between the ages of 18∼30, with normal vision and without reading disabilities.

5.2.1.3 Procedures

The participants were instructed to read sentences at their own pace, in a fashion that allowed them to fully comprehend the sentences. Sentences were visually presented segment-by-segment, shown by the cells in Table 5.3. Each trial began with a series of dash lines, whose length corresponding to the length of the words in the sentence to be read. A "+" sign appeared on the left-most edge of the first segment to signal the beginning of the sentence. The participants began by pressing the spacebar of the keyboard, causing the first segment to show up. With subsequent pressing, the following segment showed up and the previous segment just read was replaced by a set of dashes. After

[1]These participants were the same as those in Study 1

reading the sentence, the participants were queried a comprehension question. A practice session was provided before the real session, and the participants had to reach 80% accuracy before proceeding to real trials.

5.2.2 Data analysis

The data from all 28 participants were included in data analysis. The reading time (RT) measurements of the four segments of interest were analyzed: verb, complement head, and the two segments following the complement head (Complement+1, Complement+2). I performed a mixed model analysis, contrasting a model that incorporated a fixed effect of condition (4 levels) and random intercepts for subjects and items, against a model without the fixed factor. Analyses were implemented in the R statistical environment, using the *lme4* packages (Baayen et al., 2008; R Core Team, 2014). I further performed pairwise comparisons among the conditions, corrected by Tukey tests for p-values. All significant contrasts are reported below.

5.2.3 Results

The accuracy of the comprehension task was 95.03%, indicating that the participants fully comprehended the sentences and paid attention during the reading task.

At the Verb region, no significant effect of condition was found ($\chi^2(3)$=7.20,

$p< .066$); the pairwise comparisons did not show any significant difference neither. The Complement Head region did not show an effect of condition $\chi^2(3)=2.264$, $p< .52$), and no significant difference was found in the pairwise comparisons.

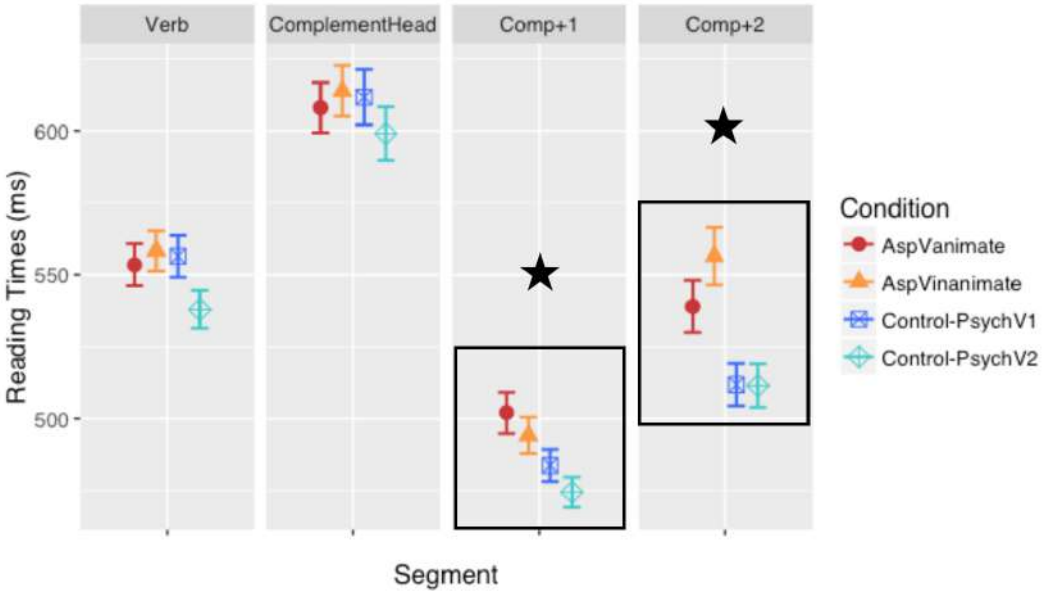
A significant effect of condition was found at the Complement+1 region ($\chi^2(3)=15.37$, $p< .0015$). At this region, the two conditions of aspectual verbs engendered significantly longer RTs than the $\text{Control}_{PsychV2}$ condition: AspVanimate > $\text{Control}_{PsychV2}$ ($b=27.58$, $p< .0015$), AspVinanimate > $\text{Control}_{PsychV2}$ ($b=19.70$, $p= .0438$).

At the Complement+2 region, a significant effect of condition was found ($\chi^2(3)= 29.04$, $p < .001$). Similar to the previous region, the pairwise comparisons showed that both conditions of aspectual verbs engendered significantly longer RTs than the two control conditions. Specifically, AspVanimate > $\text{Control}_{PsychV1}$ ($b= 27.17$, $p= .03$); AspVanimate > $\text{Control}_{PsychV2}$ ($b= 27.50$, $p= .03$); AspVinanimate > $\text{Control}_{PsychV1}$ ($b= 44.63$, $p< .001$); AspVinanimate > $\text{Control}_{PsychV2}$ ($b= 44.97$, $p< .001$). The two AspV conditions did not differ from each other, neither did the two control PsychV conditions.

Table 5.5: Study 2—Exp. 1. Results of reading times (ms), standard errors in parentheses

Condition	Verb	Complement Head	Complement+1	Complement+2
AspVanimate	553.45 (7.30)	608.01 (8.70)	502.08 (7.13)	538.99 (9.07)
AspVinanimate	558.28 (6.98)	613.86 (8.81)	494.20 (6.29)	556.45 (10.01)
$\text{Control}_{PsychV1}$	556.42 (7.26)	611.64 (9.61)	438.78 (5.55)	511.82 (7.38)
$\text{Control}_{PsychV2}$	537.93 (6.58)	598.98 (9.34)	474.50 (5.28)	511.49 (7.59)

Figure 5.3: Study 2—Exp.1. Results of reading times (ms)



Results of Experiment 1 show that AspV sentences with aspectual verbs engendered longer RTs than the psychological-verb controls at the two segments following the complement head noun. The AspVinanimate condition patterned like the AspVanimate condition, despite their difference in subject animacy. These findings are predicted by the Structured Individual Hypothesis. The longer RTs for both the AspVanimate and AspVinanimate conditions are taken to result from semantic ambiguity along multiple dimensions. Processing sentences with aspectual verbs is more costly because comprehenders have to resolve the dimension ambiguity in order to pin down an appropriate interpretation.

5.3 Experiment 2: fMRI

Experiment 2 aims to investigate whether the two configurations of aspectual verbs are processed similarly by the brain, and whether difference in subject animacy will exhibit any difference in neurological patterns.

5.3.1 Method

5.3.1.1 Materials

The stimuli contained 50 quadruples of the four conditions used in Experiment 1; a different set of 100 fillers were introduced. The whole set of stimuli

amounted to 300 sentences in total, read by each participants.

5.3.1.2 Participants

Sixteen native speakers of American English were recruited, all between the ages of 18∼30, right-handed, without reading disability nor history of neurological disorders.

5.3.1.3 Paradigm

The 300 sentences were divided equally into 10 runs (30 sentences per run). Each run lasted for 5 minutes and 33 seconds with the inclusion of device connection delay. The sentences within each run were pseudo-randomized such that no consecutive sentences were of the same condition. The order of the 10 runs was randomized for each participant. Sentences were visually presented segment-by-segment as in Experiment 2 by the E-Prime software; each segment lasted for 500 ms. At the end of the sentence, the participants received a comprehension question for 75% of the time, each remained on the screen for 4000 ms. There was a 500 ms interval interpolated between the sentence-final word and the comprehension question.

5.3.1.4 Imaging acquisition

Anatomical Measurements. The fMRI experiment was carried out on a Siemens Sonata; 3T whole body MRI scanner. Each session began with a 3-plane localizer followed by a sagittal localizer, and an inversion recovery T1 weighted scan (TE/TR=2.61/285 ms, matrix 192x192, FOV=220 mm, flip angle=70 degrees, bandwidth = 501 Hz/pix, 51 slices with 2.5mm thickness). The AC-PC (anterior and posterior commissure) line was defined by this acquisition for prescription of the anatomic T1 images and functional images in the following series.

Functional Measurements. During the task, event-related functional MRI was conducted using gradient echo echo-planar imaging (EPI) blood oxygenation level dependent (BOLD) contrast, with TE =30 ms, TR=956 ms, matrix 84 x 84, FOV=210mm, flip angle=62 degrees, bandwidth=2289 Hz/pixel, slice thickness=2.5 mm, with 321 measurements (images per slice). The scanner was set to trigger the stimulus presentation program, which enabled the image acquisition to be synchronized with the stimulus presentation.

At the end of the functional imaging, a high-resolution 3D Magnetization Prepared Rapid Gradient Echo (MPRAGE) was used to acquire sagittal images for multi-subject registration, with TE=2.77 ms, TR=2530 ms, acquisition matrix 256x256, FOV=256mm, bandwidth = 179 Hz/pix, flip angle=7 degrees, 176 slices with slice thickness=1mm. The fMRI data within subjects

was registered to this brain volume, which was then registered across subjects into a common 3D brain space by the Yale BioImage Suite software package (Papademetris et al., 2006).

5.3.2 fMRI data analysis (event-related)

All data were converted from Digital Imaging and Communication in Medicine (DICOM) format to analyze format using XMedCon (Nolfe et al., 2003). During the conversion process, the first 6 images at the beginning of each of the 10 functional runs were discarded to enable the signal to achieve steady-state equilibrium between radio frequency pulsing and relaxation, leaving 315 images per slice per run for analysis. Functional images were motion-corrected with the Statistical Parametric Mapping (SPM) 5 algorithm (www.fil.ion.ucl.ac.uk/spm/software/spm5) for three translational directions (x, y, z) and three possible rotations (pitch, yaw, roll). Trials with linear motion that had a displacement exceeding 1.5 mm or rotation exceeding 2 degrees were rejected. The data from one participant were excluded because of severe head movement. All further analyses were performed using BioImage Suite (Papademetris et al., 2006).

Individual subject data was analyzed using a General Linear Model (GLM) on each voxel in the entire brain volume with regressors specific for the tasks. As in Study 1, each sentence was segmented into two events (i.e., two regressors) in data analysis, corresponding to the two hypothesized processes

involved in the real-time comprehension of AspV sentences.

Event 1 included the onset of the subject noun phrase until the offset of the main verb. At this time window, comprehenders were hypothesized to exhaustively activate the dimension functions lexically encoded in the aspectual verb.

Event 2 included the onset of the complement noun phrase until the offset of the sentence-final word. During this time window (i.e. after encountering the complement), comprehenders were hypothesized to resolve dimension ambiguity so as to determine the dimension along which the complement is construed as an axis.

Table 5.6: Study 2—Exp. 2. Event segmentation in fMRI data analysis (cells representing the segments in stimuli presentation)

Condition	**Event 1 (Subject + Verb)**		**Event 2 (Complement ∼ Sentence-final)**			
AspVinanimate	*"Starry Night"*	*starts*	*the collection*	*of*	*impressionist*	*oil paintings.*
AspVanimate	*Van Gogh*	*started*	...	...	...	...
$\text{Control}_{PsychV1}$	*Van Gogh*	*enjoyed*	...	...	...	...
$\text{Control}_{PsychV2}$	*Van Gogh*	*loved*	...	...	...	...

The resulting beta images for each task were spatially smoothed with a 6 mm Gaussian kernel to account for variations in the location of activation across subjects. The output maps were normalized beta-maps, which were in the acquired space (2.5mm x 2.5mm x 2.5mm). To take these data into a common reference space, three registrations were calculated within the Yale BioImage Suite software package. The first registration performed a linear registration between the individual subject raw functional image and that

subjects 2D anatomical image. The 2D anatomical image was then linearly registered to the individuals 3D anatomical image. The 3D differs from the 2D in that it has a 1x1x1 mm resolution whereas the 2D z-dimension is set by slice-thickness and its x-y dimensions are set by voxel size. Finally, a non-linear registration was computed between the individual 3D anatomical image and a reference 3D image. The reference brain used was the Colin27 Brain (Holmes et al., 1998) in Montreal Neurological Institute (MNI) space (Evans et al., 1993). All three registrations were applied sequentially to the individual normalized beta-maps to bring all data into the common reference space.

Data were corrected for multiple comparisons by spatial extent of contiguous suprathresholded individual voxels at an experiment-wise $p< .05$. In a Monte Carlo simulation within the AFNI software package and using a smoothing kernel of 6mm and a connection radius of 4.33mm on 2.5mm x 2.5mm x 2.5mm voxels, it was determined that an activation volume of 183 original voxels (4953 microliters) satisfied the $p< .05$ threshold. Clusters were created for each of the subtractions. Each cluster was identified with a region label, and then associated with additional numeral labels corresponding to Brodmann areas.

5.3.3 Results

5.3.3.1 Behavioral results

Results of the comprehension task showed 88.6% accuracy. The repeated measure ANOVA showed no significant effect of condition. Mean response times to the comprehension questions were presented in Table 5.7.

Table 5.7: Study 2—Exp. 2. Response times (ms) of the comprehension questions

Conditions	Mean	SD	se
AspVanimate	1684.18	590.90	25.06
AspVinanimate	1685.85	595.22	25.57
$\text{Control}_{PsychV1}$	1722.11	582.23	25.08
$\text{Control}_{PsychV2}$	1651.49	547.17	24.09

5.3.3.2 Imaging results

At Event 1 (Subject+Verb), the AspVanimate condition engendered additional activity in left Brodmann Area (BA) 40, bilateral BA 6, 24, and BA 7, and primary sensory area, as compared to the $\text{Control}_{PsychV1}$ condition (Figure 5.4).

At Event 2 (Complement∼Sentence-final), the AspVanimate condition induced additional activity in left interior frontal cortex (LIFC), including BA 44, 45, 47, and left insula, as compared to the $\text{Control}_{PsychV2}$ condition (Figure 5.5). Other regions preferentially recruited for the AspVanimate condition over

$\text{Control}_{PsychV2}$ included bilateral BA 6, right BA 8, right IFG, and primary visual cortex. Similarly, the AspVinanimate condition preferentially recruited LIFC (BA 44, 45, 47) and left insula, as compared to the $\text{Control}_{PsychV2}$ condition (Figure 5.6). In addition, results at Event 2 revealed that the AspVanimate condition induced more activity in bilateral posterior areas (BA 17, 18, 19), bilateral BA 7, left BA 39, 31, and left BA 6, 24, as compared to the AspVinanimate condition (Figure 5.7).

Overall, the AspVanimate condition engendered more activation in BA40 (part of Wernicke's area) than the $\text{Control}_{PsychV1}$ condition at Event 1, while both the AspVanimate and AspVinanimate conditions involved LIFG and left insula at Event 2 as compared to the $\text{Control}_{PsychV2}$ condition. Besides the shared activation pattern, the brain regions recruited for the AspVanimate condition formed a superset of those recruited for the AspVinanimate condition.

Figure 5.4: Study 2—Results of Exp. 2. AspVanimate > Control-PsychV1 at Event 1

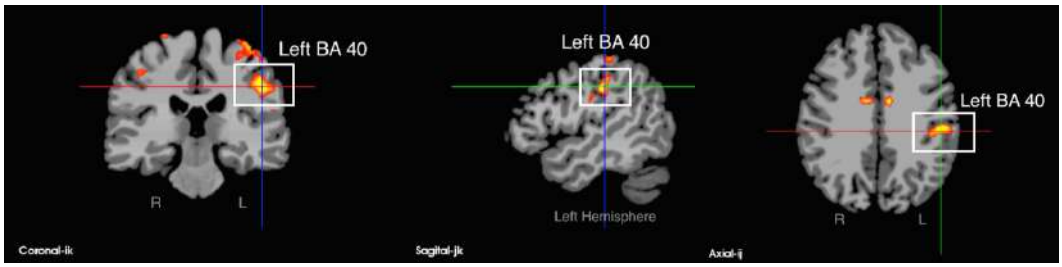


Figure 5.5: Study 2—Results of Exp. 2. AspVanimate > Control-PsychV2 at Event 2

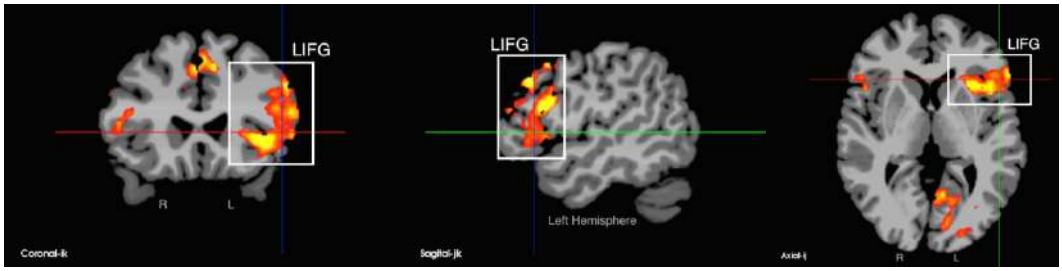


Figure 5.6: Study 2—Results of Exp. 2. AspVinanimate > Control-PsychV2 at Event 2

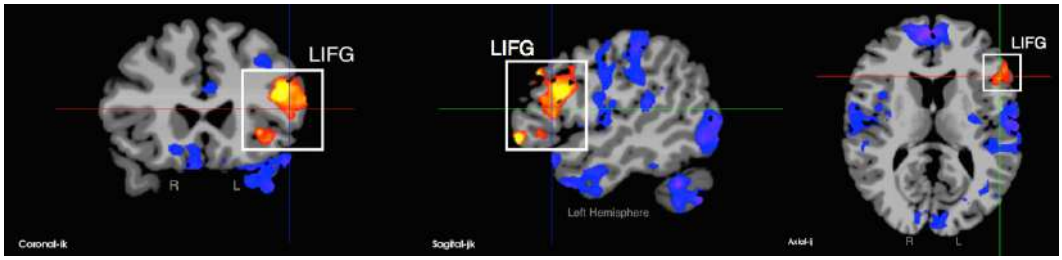
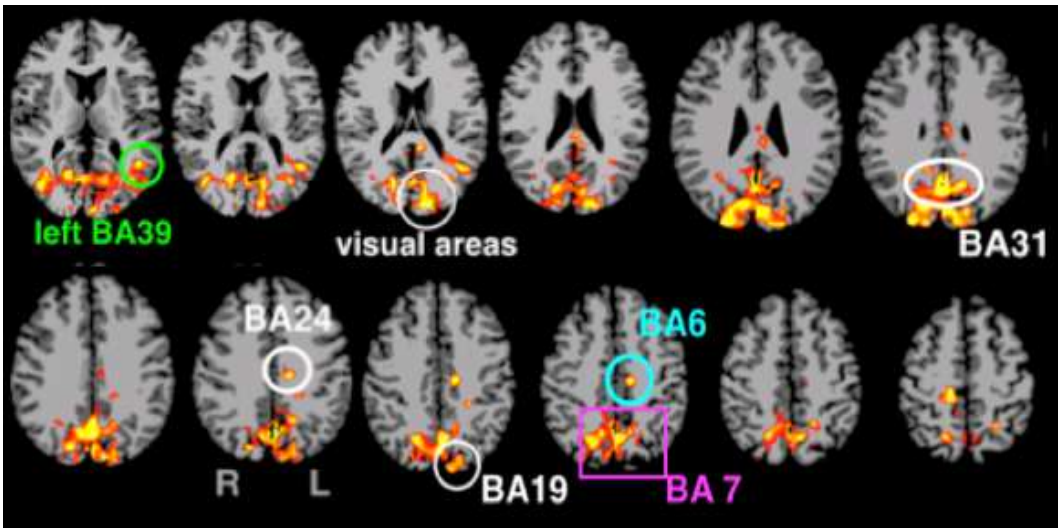


Figure 5.7: Study 2—Results of Exp. 2. AspVanimate > AspVinanimate at Event 2



5.4 Discussion

In sum, Experiment 1 (self-paced reading) showed that both the AspVanimate and AspVinanimate conditions engendered longer RTs than the two control PsychV conditions. Experiment 2 (fMRI) showed that both conditions with aspectual verbs preferentially recruited LIFG and left insula at Event 2, after the complement was encountered. Both behavioral and neurological patterns suggest that all sentences with aspectual verbs, regardless of subject animacy, are processed in a similar way. Results are consistent with the Structured Individual Hypothesis, which predicts that processing sentences with aspectual verbs involves the resolution of dimension ambiguity and hence will be more costly. The process of ambiguity resolution is likely supported by LIFC.

In addition, Experiment 2 showed that the set of brain regions recruited for the AspVanimate condition formed a superset of those recruited for the AspVinanimate condition. This superset relation reflects the available dimension readings associated with the two configurations, as shown in Figure 5.2. While the AspVinanimate configuration yields a constitutive reading along the spatial/informational/temporal/etc. dimensions, the AspVanimate configuration gives rise to an agentive reading along the eventive dimension additionally.

5.4.1 Control relation vs. Animacy

As predicted, the AspVanimate and AspVinanimate conditions patterned alike, suggesting that subject animacy alone does not determine the reading for the sentences. In fact, animacy itself is not an ideal indicator for associating semantic arguments, as it cuts across several thematic roles in the semantic representations of predicates. In sentences with the AspVanimate configuration, an animate subject referent serves as an Agent in the agentive reading, yet serves as a Theme in the constitutive reading. In the PsychV sentences, the animate subject referent serves as an Experiencer. These can be seen clearly in Table 5.8. Also, in the constitutive reading of AspV sentences, the subject referent—either animate or animate—takes a Theme role. As Levin and Hovav (2005, p.173) point out, animacy "imposes a rather coarse-grained ranking, since NPs bearing the agent, experiencer, benefactive, and recipient roles, for instance, are all typically animate." These demonstrate the insufficiency of animacy as a semantic indicator.

Table 5.8: The role of the animate subject referent in AspV/PsychV sentences

	Subject animacy	Thematic role of the subject
AspVanimate$_{-AgentiveReading}$	animate	Agent
AspVanimate$_{-ConstitutiveReading}$	animate	Theme/Stimulus (¬Agent, ¬Experiencer)
PsychV	animate	Experiencer

I suggest to replace animacy with the notion of *control relation* to capture the more fine-grained semantic representation, following Piñango (to appear).

In a situation with control asymmetry, one participant controls over anther participant (Section 5.1.3). An entity/individual with greater control ability has power over his/her own behavior and/or the resulting situation denoted by the predicate (Rozwadowska, 1989). In AspV sentences, the control relation refers to whether the participant denoted by the subject has the capacity of exercising power over the participant denoted by the complement. A greater control asymmetry between the subject and the complement denotation gives rise to an agentive reading, in which the subject referent is perceived as a controller performing some activity associated with the complement denotation. In contrast, a control symmetry (or less control asymmetry) gives rise to a constitutive reading. This notion of control relation lines up nicely with the two readings of AspV sentences, as shown in Table 5.9.

Table 5.9: Control relation and the interpretation of AspV sentences

Configuration	Example	**Control relation** (subject↔complement)	**Reading**
AspVanimate AspVinanimate	*Van Gogh began the collection.* *The printer started by paper*	Asymmetry	Agentive
AspVanimate AspVinanimate	*The little girl began the queue.* *A short postscript ends the novel.*	Symmetry	Constitutive

The parameter of control relation not only distinguishes the two readings of AspV sentences but also differentiates between AspV and PsychV sentences. The subject referent of AspV sentences in the agentive reading tends to have more control ability than the subject referent of PsychV sentences. We may employ the "*What X do is.....*" test to diagnose volitional involvement, which

is associated with control, in some action (Dowty, 1991; Culicover and Jackendoff, 2005). A set of example is given in (5.5) and tested in (5.6).

(5.5) a. John began the book.

b. The short prologue began the book.

c. John enjoyed/loved the book.

(5.6) a. AspVanimate-Agentive reading :

✓ What John *(as an actor)* did was begin the book.

b. AspVanimate-Constitutive reading :

*What John *(as a story character)* did was begin the book.

c. AspVinanimate-Constitutive reading :

*What the short prologue did was begin the book.

d. PsychV: What John did was {?enjoy/*love} the book.

The above examples show that only the subject of the AspV sentences in the agentive reading bears greater control ability unequivocally, leading to a higher control asymmetry between the subject and the complement denotations. A higher control asymmetry predicts the agentive reading of the sentences more effectively than an animate subject referent.

As hinted, the notion of control might be connectable to volitionality. The force-dynamic approach (Talmy, 1988; Croft, 2012) argues that the notion of Agent can be defined as an antagonist that has volitional intrinsic force

tendency. Dowty (1991) lists the involvement of volition as a contributing properties of his Agent Proto-Role. Jackendoff (1992, p.129) suggests the feature [±volitional] as one characteristic of Agent as "volitional Actor"[2]. Here I do not postulate a fundamental difference between the two notions, as an individual that has control over his/her behavior is typically volitional in performing the action. Using Jackendoff's test for volitionality—compatibility with *deliberately* (volitional) and *accidentally* (non-volitional), I observe that [+volition] is most compatible with the agentive reading of AspV sentences, a pattern similar to the parameter of control asymmetry (5.7).

(5.7) a. AspVanimate-Agentive reading :

✓ John *(as an actor)* began the book *deliberately.*

b. AspVanimate-Constitutive reading :

? John *(as a character)* began the book *accidentally.*

c. AspVinanimate-Constitutive reading :

*The long prologue begins the book *deliberately/accidentally.*

d. PsychV: *?*John enjoyed the book {*deliberately/accidentally*}.

The above observations and tests show that the control relation between the subject and complement denotations factors into the interpretation of AspV sentences. It is a more reliable cue that determines sentence meaning than subject animacy.

[2]For Jackendoff, Agent could be a doer of action, a volitional actor, or an extrinsic instigator.

5.5 Conclusion

In this chapter, I presented two experiments that validate the SIH's that all sentences with aspectual verbs pattern alike due to exhaustive dimension composition and dimension ambiguity. The behavioral and neurological findings suggest that sentences with aspectual verbs engender a similar processing profile regardless of subject animacy. Moreover, the fMRI experiment shows that the AspV sentences with an animacy subject referent induced additional brain activity than those with an inanimate subject referent. This neurological pattern reflects the superset relation between the AspVanimate and AspVinanimate configurations in terms of their available dimension readings. I suggest that the ultimate reading of AspV sentences is influenced by (a) the control relation between the subject and the complement denotations, and (b) the choice of dimension *in combination*. The notion of control relation better predicts the agentive versus constitutive reading of AspV sentences than subject animacy.

Study 1 (Chapter 4) & Study 2 (Chapter 5) show the effect of dimension composition of the verb's functions and the complement denotation. By hypothesis, this dimension composition takes place in all cases regardless of context (see Chapter 3). However, it remains a question regarding whether and how context comes into play in the real-time comprehension of AspV sen-

tences. In the next chapter, I will present a eye-movement study that addresses this question.

Chapter 6

Real-time comprehension of aspectual-verb sentences in context: An eye-tracking experiment (Study 3)

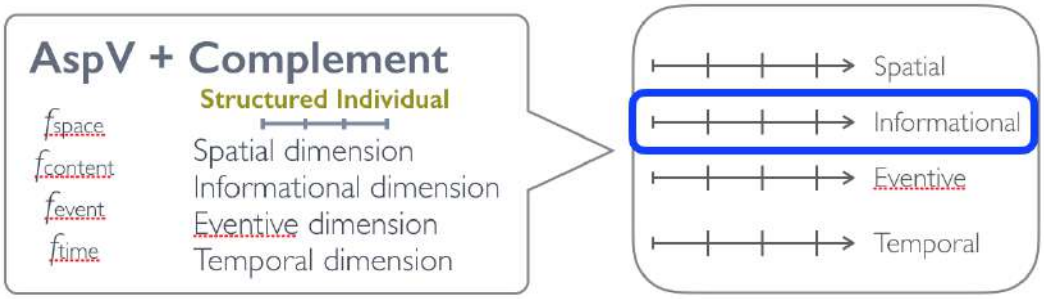
6.1 Introduction

Study 1 (Chapter 4) and Study 2 (Chapter 5) show that processing AspV sentences engenders longer reading times and preferential brain activity than the control sentences with psychological verbs (PsychVs), consistent with the Structured Individual Hypothesis (SIH). However, in Study 1 & 2, the sen-

tences are tested without context. This chapter presents an eye-movement study that investigates the role of context in the processing of aspectual verbs. To this end, I examine the processing profiles of AspV sentences following neutral versus biasing contexts.

It is indicated in Chapter 3 (Section 3.3.1) that regardless of context, the composition of the verb's dimension functions and the complement takes place in all AspV sentences. The composition gives rise to multiple dimension representations (e.g. spatial, informative, eventive), and comprehenders determine a particular dimension interpretation according to the context in which the AspV sentence appears. On the SIH, the role of context is to constrain (not pre-determine) the dimension representations *after* they are composed from the verb and the complement (Figure 6.1). This means that the cost associated with lexical retrieval and dimension composition for AspV sentences cannot be eliminated by context.

Figure 6.1: The dimension composition of an aspectual verb and the complement, followed by contextual constraining



On the other hand, another source of cost associated with AspV sentences is the resolution of dimension ambiguity (see Chapter 3). One would expect that a prior context biasing towards one particular reading (either agentive or constitutive) would guide comprehenders towards the contextually-supported reading for the AspV sentence in question. Consequently, the cost associated with ambiguity resolution would be attenuated. In this case, AspV sentences in a biasing context are expected to engender less processing cost than those in the neutral context which is compatible with both readings.

One point should be noticed here: the effect of ambiguity resolution based on context might be influenced by the availability of the alternative interpretations. For AspV sentences, especially sentences with an animate-denoting subject (the AspVanimate configuration), it is not clear whether the agentive and constitutive readings are equally available in the neutral context. As indicated in Chapter 5, one factor that influences the choice between the agentive versus constitutive reading is the *control relation between the participants* in the subject and the object positions. On the assumption of incremental processing, an animate subject referent could lead to an initial bias during real-time comprehension: because animate entities (as compared to inanimate ones) are perceived as controllers more often, an animate subject referent will increase the control asymmetry between the participants at least temporarily. This greater control asymmetry gives rise to an agentive reading along the eventive dimension, such that the subject denotation is conceived as a

controller (agent) performing some activity associated with the complement denotation. If an animate subject referent carries an inherent preference for greater control asymmetry, comprehenders may adopt the agentive reading when they encounter AspVanimate sentences in a neutral context. If comprehenders rely heavily on this preference to deal with the ambiguity between the agentive vs. constitutive reading, then it is possible that we will *not* observe more cost for AspVanimate sentences in a neutral context. In other words, the control asymmetry guided by the animate subject referent may mask the effect of ambiguity brought about by the neutral context. Alternatively, if the inherent preference due to the animate-denoting subject weighs significantly less than contextual bias (or if the context bias takes effect faster enough to overwrite such preference), we would be able to observe more cost in AspVanimate sentences following a neutral context as compared to a disambiguating context.

To investigate these issues, I conducted an eye-tracking experiment, asking participants to read sentences following a neutral context versus a biasing context that supports either the agentive or constitutive reading. AspV sentences are contrasted with psychological-verb (PsychV) sentences as controls. In any case, the SIH predicts that AspV sentences, which involve dimension composition and ambiguity resolution, should require more processing effort than the control PsychV sentences when the context remains constant between the two verb types.

6.2 Method

The method follows Katsika et al.'s (2012) eye-tracking experiment, because it is the first study that singles out aspectual verbs from psychological verbs and shows additional cost for aspectual verbs only. I expect to replicate their findings of the verb type effect, as in Study 1 and 2. Furthermore, I intend to compare the processing patterns of AspV sentences with and without context.

6.2.1 Participants

Fifty-one native speakers of American English were recruited (36 females). All were between the ages of 18∼30, with normal or corrected vision and without any reading disabilities. The protocols of this study were approved by the Yale University Human Subjects Committee. Participants gave informed consent prior to the experiment and received monetary compensation.

6.2.2 Apparatus

Binocular eye-movements were recorded by an Eye-Link 1000 Plus eye-tracker, at the 500Hz sampling rate. The participants were instructed to place their head on a head-mount; the distance between the head-mount and the the screen for stimuli presentation was about 68 cm. Sentences were displayed in a

monospace font (Consolas), with 3.39 characters per degree of visual angle. All target sentences fitted in one line, left aligned horizontally with the left margin as 85 pixels and vertically centered on the screen. The context sentences were displayed in one or two lines on the screen depending on their lengths; the left margin remained the same as that of the target sentences.

6.2.3 Materials

Thirty sets of passages were employed, each consisting of a context sentence and a target sentence. Each target sentence contained either an aspectual verb or a psychological verb. The target sentences were adapted from Study 2, modified minimally to bridge the content of the context sentence. Each AspV target sentence was paired with three types of context: (1) a context biasing towards the agentive reading along the eventive dimension (**eventBiasing-AspV**), (2) a context biasing towards the constitutive reading along the informational or spatial dimension (**constBiasing-AspV**), and (3) a neutral context that allowed both the agentive and constitutive readings (**Neutral-AspV**). In the eventBiasing condition, the context sentence mentioned some action performed by the subject as a controller (taking the Agent role). In the constBiasing condition, an entity referring to a structured individual was made salient and no action was mentioned. In the Neutral condition, the context sentence was manipulated to allow both the agentive reading and the constitutive reading. In

data analysis (Section 6.2.5), I examined the effect of context as a fixed factor with 3 levels of context type: constBiasing, eventBiasing, Neutral.

On the other hand, each PsychV sentence was paired with the eventive-biasing or neutral context (**eventBiasing-PsychV**, **Neutral-PsychV**). PsychV sentences were not paired with the constitutive-biasing context because they do not engender a constitutive reading (such that the subject is conceived as a subpart of the entity denoted by the complement). The two PsychV conditions here contained both verbs like *enjoy* and *endure*, which are considered as coercion verbs on the Type-Shifting account, and others like *love* and *hate*. Recall that these two sets of psychological verbs were referred to as EnjoyingVs and LovingVs in Study 1 (Chapter 4) and Study 2 (Chapter 5). As mentioned there, the Structure Individual Hypothesis does not distinguish these sub-types of psychological verbs, and they in fact patterned alike in real-time comprehension in Study 1 & 2. In the present experiment, the PsychV conditions mainly served as the baseline to compare with the AspV conditions. The emotion polarity (positive/negative) of psychological verbs was balanced as much as possible across the stimuli.

In addition, fifty filler sentences were introduced. Among these, half of the target sentences were grammatical and half were ungrammatical; each grammatical/ungrammatical filler condition contained half AspVs and half PsychVs. Thirty items of the fillers were with the constBiasing context to counterbalance the context types across the stimuli.

The whole set of stimuli (30 sets of 5 conditions plus 50 fillers) amounted to 200 sentences in total. An example set of the five critical conditions was provided in Table 6.1. The critical verbs used are provided in Table 6.2, with their frequencies as appeared in the stimuli shown in the parentheses.

Table 6.1: Study 3—Conditions and sample sentences

Condition	**Context Sentence**	**Target Sentence**
eventBiasing-AspV (EA)	Musicians often record their pieces for compilation or memorial albums.	Dave Brubeck **started** this CD of classic Jazz hits.
constBiasing-AspV (CA)	Kevin owns numerous CDs by different Jazz musicians.	Dave Brubeck **started** this CD of classic Jazz hits.
Neutral-AspV (NeuA)	Many musicians have music libraries that contain tons of albums.	Dave Brubeck **started** this CD of classic Jazz hits.
eventBiasing-PsychV (EP)	Musicians often record their pieces for compilation or memorial albums.	Dave Brubeck **loved** this CD of classic Jazz hits.
Neutral-PsychV (NeuP)	Many musicians have music libraries that contain tons of albums.	Dave Brubeck **loved** this CD of classic Jazz hits.

Table 6.2: Study 3—Verbs and their frequencies used in the 5 critical conditions

AspV (30)	*start* (8), *finish* (5), *end* (4), *begin* (9), *continue* (2), *complete* (2)
PsychV (30)	LovingV (15): *love* (2), *detest* (5), *respect* (1),*like* (4), *dislike* (2), *hate* (1)
	EnjoyingV (15): *tolerate* (3), *enjoy* (4), *prefer* (4), *favor* (2),*endure* (2)

The stimuli were pseudo-randomized for each participant. To this end, four master scripts for the stimuli were created; each master script consisted of 30 blocks of the 200 items, each block containing 6 or 7 items of different conditions with different item sets (e.g. the five conditions in Table 1 formed

a set). I then used these four master scripts to randomly generate distinctive subject scripts presented to the participants. For each subject script, the block order was different and the order of the items within each block was different as well.

Furthermore, to ensure full attention from the participants, comprehension questions were given approximately 1/3 of the time (68 trials out of 200) presented immediately after the target sentence. The questions were interspersed randomly throughout the trials and distributed evenly across the conditions; the correct answers were half "Yes" and half "No." To avoid strategic reading, half of the comprehension questions queried the content of the context sentences and the other half queried the content of the target sentences.

6.2.4 Procedures

Participants were instructed to place their head on a head-mount while seated in front of the display screen. The whole eye-tracking experiment was run in three sub-sessions with breaks in between. Each sub-session began with a screen presenting the task instruction, followed by the eye-tracking calibration with a series of 13 fixation targets across the display.

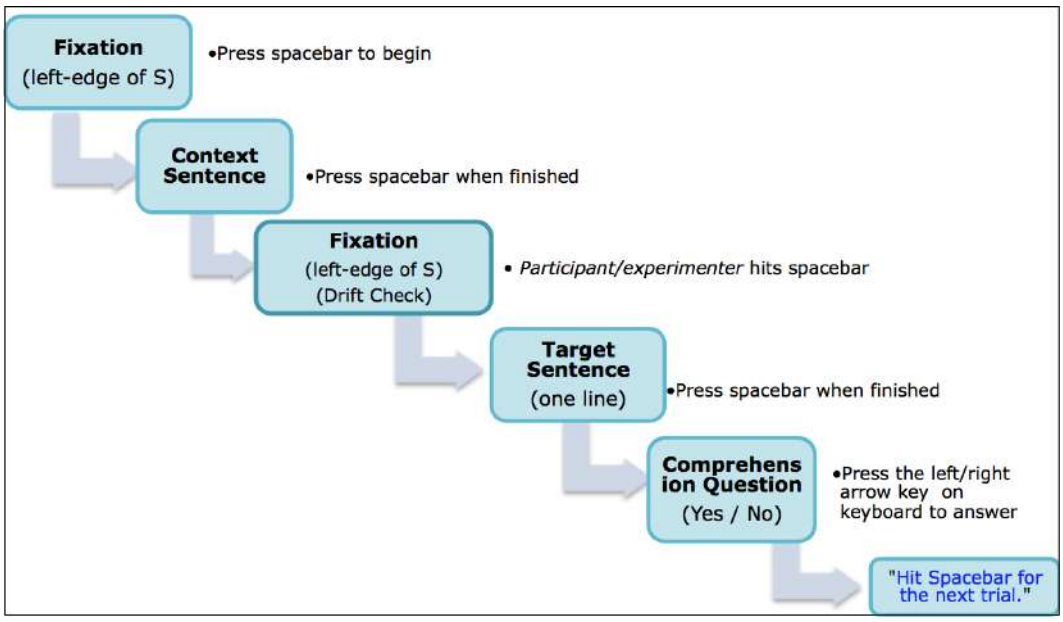
Each trial began with a target dot (presented as a fake drift check) at the left-edge of the screen to signal the starting point of the context sentence. The participants were instructed to fixate on the target dot and pressed the

spacebar, which brought up the context sentence. After they finished reading the context sentence, the participants hit the spacebar. Then another fixation target dot, manipulated as a drift check, appeared at the left-edge of the screen; again, the participants fixated on the dot and pressed spacebar to bring up the target sentence. They pressed the spacebar when they finished reading the target sentence.

For 1/3 of the trials, a comprehension question was shown immediately after the target sentence in those trials. The participant pressed the left arrow key ($\Leftarrow$) for "Yes" and the right arrow key ($\Rightarrow$) for "No." The *Yes* and *No* options were provided on the screen at the positions corresponding to the arrow keys on the keyboard (i.e., "Yes" on the left; "No" on the right). Once the response to the question was made, the display screen presented a page showing "*Hit spacebar for the next trial.*", employed as a transition break between trials. For trials which did not have a comprehension question, the display screen went directly to this transition page.

Prior to the real session, a practice session with six trials was given in order to familiarize the participants with the experimental procedures. The whole trial procedure is provided in Figure 6.2.

Figure 6.2: Study 3—Eye-tracking experiment. Trial procedure



6.2.5 Data analysis

The data from one participant were excluded because the fixations from this participant were outside of the predetermined vertical boundaries of the interest areas (122 pixels above and 100 pixels below the baseline, i.e., the bottom line of the English letters) for more than 25% of the trials.

Fixations shorter than 50 ms were merged with nearby fixations (*cf.* Carpenter and Just, 1983; Katsika et al., 2012). Prior to analyses, I removed data from critical items by the following criteria: (a) trials which contained more than three blinks (2.947% trials removed), (b) trials whose duration was longer than 10 seconds (0.227% trials removed), (c) trials whose first fixation did not land at the region encompassing the subject noun phrase (Region 1 as defined below) (4.548% trials removed), and (d) trials which were not properly recorded due to experimental errors (1.307% trials removed). By these criteria, 9.029% of the critical trials were removed in total. While the eye-tracking recording was binocular, data from the right-eye were used as default. For four participants whose right-eye data appeared messy by visual inspection due to recording difficulties, the left-eye data were used.

For the analyses of eye-movement measurements, each target sentence was segmented into five regions, following Katsika et al.'s (2012) region definitions. *Region 1* contained the subject noun phrase; *Region 2* contained the critical main verb up to the word prior to the complement head. *Region 3* contained

the complement head and the following word. *Region 4* contained the two words following Region 3; if both these two words were not content words, we included the third word following Region 3 to ensure that the region had at least one content word. *Region 5* included the remainder of the sentence. An example of region segmentation is shown in Table 6.3.

Table 6.3: Study 3—Example of region segmentation

Verb Type	Reg. 1	**Reg. 2**	**Reg. 3**	**Reg. 4**	Reg. 5
AspV	Dave Brubeck	*started* this	CD of	classic Jazz	hits.
PsychV	Dave Brubeck	*loved* this	CD of	classic Jazz	hits.

I focused on the three regions of interest here: **Region 2** (main verb ∼ the word prior to the complement head), because comprehenders are hypothesized to retrieve all dimension functions encoded in aspectual verbs; **Region 3** (Complement head & the following word), where the available dimensions were retrieved from the complement head. **Region 4** (two words after Region 3) was included to avoid missing any lasting effect associated with the process of ambiguity resolution.

For each region of interest, four measurement variables were examined, following Katsika et al. (2012). (i) *First-pass duration* is the sum of all first-run fixations within a target region, beginning from the first fixation in the region until the eyes leave it. (ii) *First-pass regression* evaluates whether the sequence of first-pass fixations in a region ends in a backward glance to an earlier region; this is coded dichotomously as 0/1— 1 if the comprehenders

look back to an earlier part of the sentence after the first fixation in the region in question. (iii) *Regression-path duration* includes the sum of all fixation durations from the first fixation in a region until the eyes exist the region. (iv) *Second-pass duration* is the sum of all fixation durations in a region that are not the first-pass fixations.

Eye-movement data were analyzed by mixed effects models, carried out in the R statistical environment (R Core Team, 2015) using the *lme4* packages. For the measurements with continuous responses (regression-path duration, second-pass duration), linear mixed models were used. For the first-pass regressions with dichotomous responses, logistic mixed effects models were used. Two sets of analyses were implemented to address two different research questions. The first set targeted the conditions with aspectual verbs only, paired with three different context types (Neutral-AspV, eventBiasing-AspV, constBiasing-AspV). This aimed to evaluate the effect of context type on the comprehension of AspV sentences. An effect of context was assessed by contrasting a model that incorporated a fixed factor of context type against a model without it.

The second set of analyses targeted the effect of verb type, context type, and their interaction in sentence comprehension, as a 2 x 2 design: Verb Type (AspV, PsychV) x Context Type (Neutral, eventBiasing). I examined the interaction effect of verb type and context type by contrasting the full model that incorporated these two fixed factors and their interaction against the

additive model that dropped the interaction term. The effect of verb type was assessed by contrasting an additive model that incorporated fixed factors of both context type and verb type against a model that dropped the fixed factor of verb type. Likewise, the effect of context type was assessed by contrasting the additive model against a model that drops the fixed factor of context type. In both sets of analyses, all models included random intercepts for subjects and items (Baayen, 2008)[1].

In addition, Bonferroni correction was performed when assessing the significance of any effect. The critical p-value of 0.05 was divided by 12 (3 regions of interests x 4 measurement variables), thresholding at p=0.0042. All signifiant results are reported below.

6.3 Results

Comprehension question accuracy (n=50) was 92.79%; all participant reached 80% accuracy.[2] This suggests that the participants were paying attention during the reading task.

Now let's turn to the eye-movement results. The first set of analyses on AspV-only conditions (Neutral-AspV, eventBiasing-AspV, constBiasing-

[1]Models with random slopes failed to converge for some measurement variables and regions; hence the random intercept models were used for consistent across all regions of interest and variables.

[2]The overall accuracy for all 51 participants was 92.82%; all participants reached 80% accuracy.

AspV) yielded no significant results at any region in any measurement.

The second set of analyses on Verb Type (AspV, PsychV) x Context Type (Neutral, eventBiasing) revealed a Verb Type effect in all regions of interest, such that sentences with aspectual verbs engendered longer reading times (RTs) than sentences with psychological verbs in later measurements. I did not find an effect of context, nor a Context Type x Verb Type interaction.

Specifically, Region 2 showed a significant main effect of Verb Type, such that AspVs > PsychVs in second-pass duration ($\chi^2(1)= 29.796$, $p< .001$).

Region 3 also showed a main effect of Verb Type such that AspV > PsychV, found in first-pass regression ($\chi^2(1)= 31.168$, $p< .001$) and second-pass duration ($\chi^2(1)= 19.996$, $p< .001$). However, the first-pass duration showed a reverse Verb Type effect, such that PsychV > AspV ($\chi^2(1)=10.563$, $p= 0.001$).[3]

At Region 4, the Verb Type effect such that AspV > PsychV appeared as well, found in regression-path duration ($\chi^2(1)= 9.824$, $p= .0017$). No effect of Context Type was found, neither was there an interaction between Verb Type and Context Type. The means and standard errors are provided in Table 6.4 and visually presented in Figure 6.3∼6.6.

[3]Post-hoc pairwise comparisons with Tukey corrections between the conditions suggested that this was driven by Neutral-PsychV, which engendered longer RTs than Neutral-AspV ($p< .0048$) and eventBiasing-AspV ($p< .0036$). The comparison between Neutral-PsychV and Neutral-AspV did not survive Bonferroni correction thresholding at $p< .0042$ however.

Table 6.4: Study 3—Means and standard errors (in parentheses) of RTs

Condition	**Region**	**First-Pass Duration**	**First-Pass Regression**	**Regression-Path Duration**	**Second-Pass Duration**
Neutral-AspV		339.744 (5.815)	0.062 (0.007)	386.921 (7.835)	266.370 (9.481)
eventBiasing-AspV		336.686 (5.688)	0.069 (0.007)	379.333 (7.312)	271.968 (9.615)
constBiasing-AspV	Reg. 2	340.903 (6.007)	0.089 (0.008)	397.476 (7.836)	274.122 (10.292)
Neutral-PsychV		344.066 (5.687)	0.059 (0.006)	378.229 (6.870)	220.603 (8.530)
evetBiasing-PsychV		338.042 (5.405)	0.052 (0.006)	372.394 (7.307)	229.052 (8.278)
Neutral-AspV		319.720 (5.795)	0.248 (0.012)	466.178 (10.684)	247.583 (9.292)
eventBiasing-AspV		319.723 (5.405)	0.258 (0.012)	471.758 (10.429)	249.826 (9.158)
constBiasing-AspV	Reg. 3	318.980 (5.813)	0.239 (0.012)	481.183 (12.034)	249.977 (9.277)
Neutral-PsychV		345.193 (6.534)	0.191 (0.011)	460.524 (11.113)	203.170 (8.195)
eventBiasing-PsychV		328.422 (5.641)	0.193 (0.011)	444.865 (9.814)	225.031 (8.849)
Neutral-AspV		408.454 (8.217)	0.438 (0.013)	855.742 (20.859)	246.299 (10.524)
eventBiasing-AspV		430.356 (9.630)	0.424 (0.013)	888.685 (22.247)	250.469 (11.078)
constBiasing-AspV	Reg. 4	415.694 (8.827)	0.414 (0.013)	844.533 (21.880)	256.864 (10.563)
Neutral-PsychV		435.589 (8.919)	0.397 (0.013)	818.354 (20.212)	228.609 (10.136)
eventBiasing-PsychV		409.349 (8.035)	0.415 (0.013)	817.771 (19.542)	228.383 (9.616)

In summary, the eye-movement results showed a main effect of Verb Type (AspV > PsychV) at Region 2∼4. The effect appeared in first-pass regression, regression-path duration, and second-pass duration. I did not find an effect of Context Type, nor an interaction between Verb Type and Context Type. In the next section, I discuss our interpretations for these results.

Figure 6.3: Study 3—First-pass durations. (Means with ±1 SE)

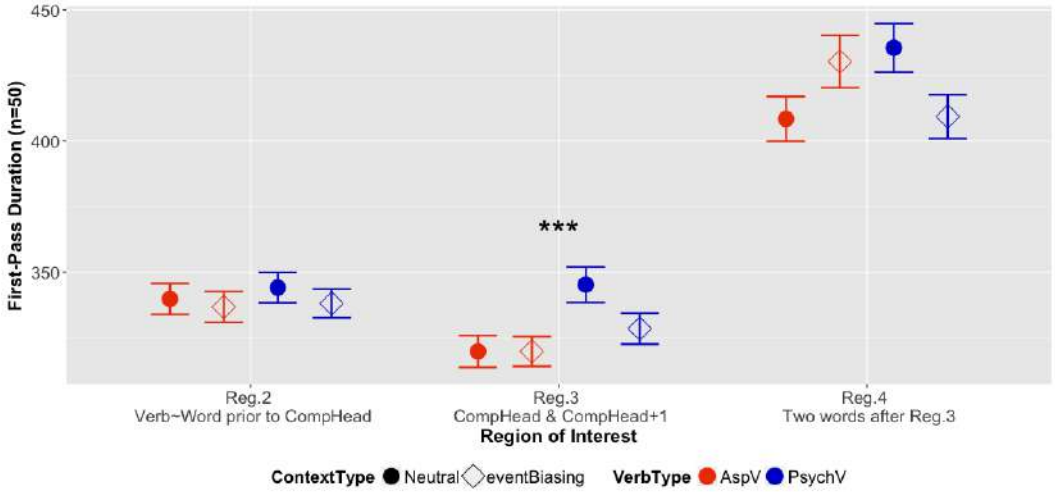


Figure 6.4: Study 3—First-pass regressions. (Means with ±1 SE)

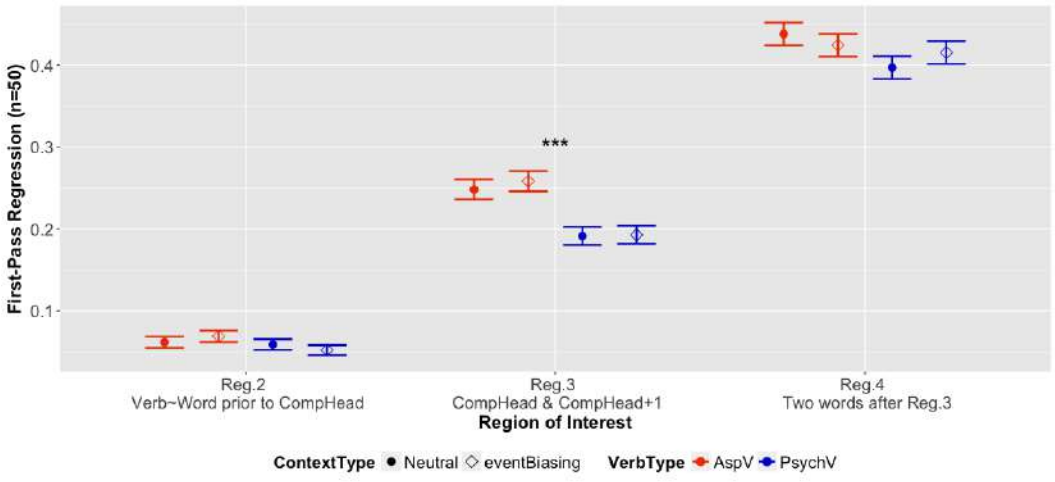


Figure 6.5: Study 3—Regression-path durations. (Means with ±1 SE)

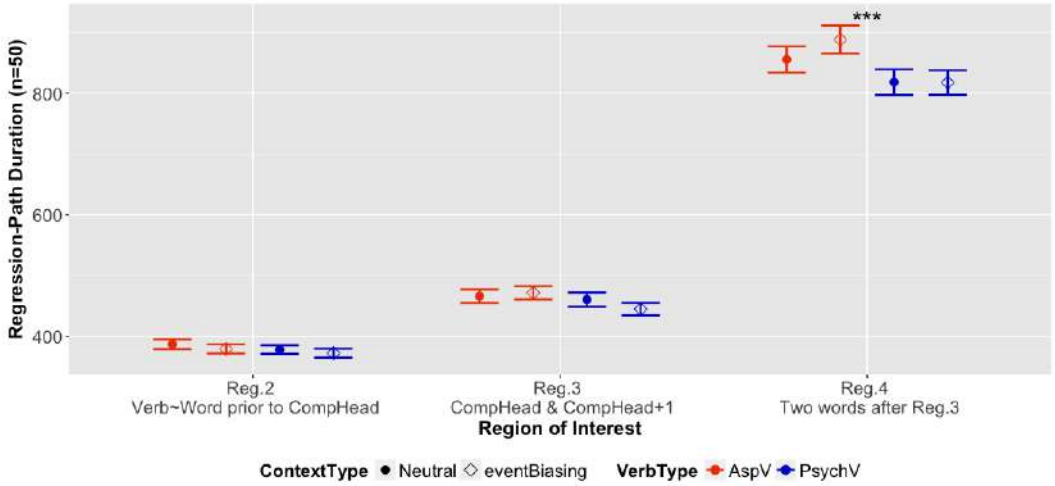
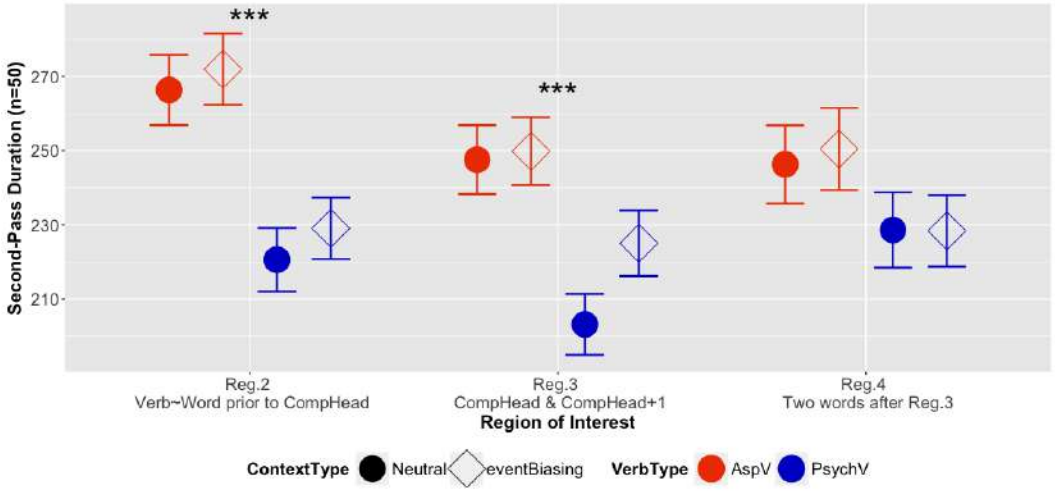


Figure 6.6: Study 3—Second-pass durations. (Means with ±1 SE)



6.4 Discussion

To begin with, the results obtained here replicate Katsika et al.'s (2012) study, which shows longer RTs for aspectual verbs than psychological verbs. Notice that the current study uses the same definitions for the regions of interest as in that study. In Katsika et al.'s study, the contrast between the two verb types appeared in second-pass times in Region 2 (main verb∼the word prior to the complement head), and in first-pass regressions in Region 3 (the complement head and the following word). With the same region segmentation, the effect of Verb Type was found in exactly the same measurements and regions: second-pass times in Region 2, and first-pass regressions in Region 3. Additionally, the verb type effect was observed in second-pass durations in Region 3, and regression-path durations in Region 4 as well.

The effect for Verb Type reported in the current study appeared in all regions of interest. Consistent with the SIH's prediction, comprehenders spent more time reading sentences with aspectual verbs than sentences with psychological verbs. Recall that on the SIH, the additional cost associated with aspectual verbs comes from two sources: (a) exhaustive activation of the dimension functions encoded in the verbs along with the dimension composition with the complement, and (b) resolution of dimension ambiguity among multiple dimension representations. This interpretation marks one crucial difference

between Katsika et al.'s interpretation and the current one: while Katsika et al. attribute the additional cost associated with AspV sentences to the type-shifting operation that coerces the entity-denoting complement into an event, the current hypothesis attributes the cost to the dimension composition and ambiguity resolution.

The verb type effect found in first-pass durations in Region 3 (Figure 6.3) exhibited an opposite pattern, such that PsychV sentences engendered longer RTs than AspV sentences, driven by Neutral-PsychV (> Neutral-AspV, and > eventBiasing-AspV) This not predicted by the SIH, and I do not have a clear answer at this point and leave it for future investigation.

The exhaustive retrieval of the dimension functions encoded in AspVs is associated with additional processing effort in DiNardo's (2015) ERP study. It shows that regardless of subject animacy, AspV sentences (without context) induced larger N400 after the onset of the verb and P500 after the onset of the complement head, as compared to psychological verbs. The N400 effect for aspectual verbs is claimed to reflect the retrieval of all lexical functions in the verb, while the P500 effect—argued to equate P600—reflects the process of ambiguity resolution among multiple dimension interpretations. That is, for any given sentence with an aspectual verb, the lexical functions in the verb are exhaustively retrieved, hence the N400 effect for aspectual verbs.

The issue in question is how does context influence the real-time comprehension of sentences with underspecified meaning, as in AspV sentences.

It turns out that the reading times in the three regions examined here (Region 2∼4) did not show an effect of context type (i.e., eventBiasing-AspV, constBiasing-AspV, and Neutral-AspV did not differ). This is not expected and seems puzzling. Below I lay out my speculation for the absence of context effect.

First of all, it is indicated in Section 6.1 that the dimension composition between the verb's functions and the complement denotation takes place regardless of context (Figure 6.1). The cost associated with this process would not be affected by context, whose role is hypothesized to constrain the plausible dimension interpretation after the dimension composition. It is likely that the processing profiles of the three AspV conditions with different context types all reflect this process of dimension composition, hence the cost is comparable in all contexts.

Second, the factor of context might take effect slower than expected and did not accrue in the three regions examined here. A possible way to test this explanation is to make the target sentences longer and examine the segments downstream in the sentences. If this interpretation is on the right track, then we should find longer RTs for the neutral-context condition, as compared to the biasing-context conditions, in later regions of AspV sentences.

The factor of ambiguity resolution is less clear based on the current findings of the AspV conditions with the three context types. On one hand, the two biasing-context conditions patterned alike as expected. If the cost as-

sociated with AspV sentences results from ambiguity resolution, one would expect that a biasing context with disambiguating information could lessen the effort of ambiguity resolution because it guides comprehenders towards the contextually-supported interpretation. With the eventBiasing context, the agentive reading stands out; with the constBiasing context, the constitutive reading stands out. With the help of informative context, the cost of processing the target AspV sentences in the eventBiasing and constBiasing conditions is attenuated, resulting in no difference between the two.

On the other hand, the Neutral-AspV condition did not engender more cost than the two biasing conditions. I suspect that the absence of context effect has to do with meaning dominance of the AspVanimate configuration employed in the current experiment. That is, the agentive vs. constitutive reading may not be equally available to the same degree for AspV sentences with an animate subject referent. In what follows, I tease apart the contrast between the eventBiasing-AspV vs. Neutral-AspV condition and the contrast between the constBiasing-AspV vs. Neutral-AspV condition respectively.

6.4.1 Neutral-AspV vs. eventBiasing-AspV

Why did the Neutral condition not show more processing cost than the eventBiasing condition? Notice that our sentences all contained an animate subject referent. I have argued in Chapter 5 (Study 2) that the ultimate

reading of AspV sentences is determined by two factors: (a) the control relation between the subject referent and the complement denotation, in addition to (b) contextual information for determining the exact dimension. While the context could help disambiguate the reading, the control relation between the participants of a given sentence has impact on sentence comprehension in real-time still.

Although the AspV sentences with an animate subject referent (AspVanimate) in the neutral context allow both the agentive and the constitutive readings, the two readings may not equally probable. An animate subject referent in isolation is more frequently conceived as a controller (taking an Agent role), carrying an inherent bias towards a higher control asymmetry that gives rise to the agentive reading (the agentive subject performing some action associated with the complement). That is, for AspVanimate sentences, the agentive reading is more dominant than the constitutive reading in the absence of conflicting contextual constraints.[4] Suppose that the frequency of alternative readings is associated with the degree and/or speed of activation, and that the processor evaluates the alternatives in the order of their activation levels (Braze, 1999; Jackendoff, 2007; Seidenberg et al., 1982; Tanenhaus et al., 1979). In this case, the agentive reading will be adopted for AspVanimate sentences because this reading is activated to a greater degree due to

[4]To further clarify this issue, it would be useful to examine the ratio of animate to inanimate subject referent with AspV sentences in the corpus.

the animate-denoting subject and because it is compatible with the neutral context (Rayner and Frazier, 1989).

In other words, the advantage of meaning-frequency prioritizes the agentive reading in the Neutral-AspV condition, given that an animate subject referent increases the control asymmetry which gives rise to the agentive reading. As a result, the Neutral condition is to some extent similar to the eventBiasing condition in that the agentive reading is more readily available to comprehenders in both cases. For the Neutral condition, the advantage of the agentive reading results from the meaning-frequency associated with the animate subject referent that increases the likelihood of higher control asymmetry; the cost of ambiguity resolution due to meaning uncertainty is thus less than expected. For the eventBiasing condition, the advantage of the agentive reading could result from either the same meaning-frequency or the biasing contextual information, or both. Consequently, the Neutral condition and the eventBiasing condition did not differ significantly. As Rayner and Frazier (1989, p. 786) indicate, "a biased word and an unambiguous word are indistinguishable when the dominant (or only) meaning of the word occurs frequently and when it is easily integrated with prior context. A strongly biased prior context can facilitate integration in either case."

6.4.2 Neutral-AspV vs. constBiasing-AspV

The finding that the constBiasing and Neutral-AspV conditions patterned alike raises two issues. On one hand, one would expect the constBiasing-AspV condition to show less cost than the Neutral-AspV condition because the disambiguating contextual information in the constBiasing condition should attenuate the cost associated with ambiguity resolution. Yet no difference was found, and the reason could be the same as above. The inherent meaning bias rooted in the animate subject referent guides the processor towards the agentive reading in the Neutral condition. Therefore, the cost associated with ambiguity resolution in the Neutral condition is attenuated, becoming comparable to the biasing condition. That is, the pattern {eventBiasing-AspV = constBiasing-AspV = Neutral-AspV} is due to the inherent bias of the Asp-Vanimate configuration towards the agentive reading.

On the other hand, one might expect the constBiasing-AspV condition, which facilitates the subordinate constitutive reading, to show *more* cost than the Neutral-AspV condition due to the *subordinate bias effect* reported in literature (Binder and Rayner, 1998; Duffy et al., 1988; Pacht and Rayner, 1993; Rayner and Frazier, 1989; Sereno et al., 2006). This effect refers to the phenomenon that an unbalanced ambiguous word (i.e., a word that has a more frequent dominant meaning and a less frequent subordinate meaning) in a context supporting the subordinate meaning engendered *more* cost, as compared

to an unambiguous control item that matches the word-form frequency, even in a strongly biasing context. This comes with the assumption that an unbalanced ambiguous word in a neutral context yields the dominant meaning and thus patterns like an unambiguous word. The ambiguity of syntactic structure also shows similar subordinate bias effect. For example, Van Gompel et al.'s (2001) study on the NP/VP-attachment ambiguity reveals that the disambiguated condition towards a less frequent interpretation is more costly than the ambiguous condition.

Duffy et al.'s (1988) Reordered Access Model contends that all meanings of an ambiguous words are exhaustively accessed; yet the availability of each word meaning are ordered by frequency and context. According to this model, the subordinate bias effect arises because the context supports the subordinate meaning while frequency supports the dominant meaning. Due to the biasing contextual information, the subordinate meaning is available earlier than usual (as in the case of neutral context in which the dominant meaning would be available first). The equal degree of availability of the two meanings causes processing difficulty due to competition. On the other hand, for Rayner and Frazier (1989), who argue for the Integration Model, lexical meanings are not always exhaustively accessed; once the processor successfully integrates one meaning with the context, the access to the other meaning is excluded.[5]

[5]This could explain the no difference between the Neutral-AspV and eventBiasing-AspV conditions, but I assume that the order of evaluation is at the activation level while assuming exhaustive lexical retrieval, following the constraint-based approach (Jackendoff, 2007; MacDonald et al., 1994a,b; Seidenberg et al., 1982; Tanenhaus et al., 1979).

According to this model, the subordinate bias effect arises because integrating the dominant meaning with the context supporting the subordinate meaning yields no coherent reading; hence "the processor must wait for the subordinate meaning to become available from the lexicon before proceeding with sentence integration."

If there is a subordinate effect for unbalanced ambiguity, the constBiasing condition, in which the context biases towards the subordinate reading, should be more costly than the eventBiasing condition and the Neutral condition (the latter two patterned alike). However, no difference between these conditions was found. I conjecture that the dis-preference for the constitutive reading due to an animate subject referent and the contextual bias towards that reading balance out. That is, the factor of *meaning-domiance* results from the inherent control asymmetry between the participants and the factor of *contextual bias* cancel out each other's impact, hence eliminating the potential difference between the constBiasing and the Neutral condition. This can be verified in the future by contrasting AspV sentences with an animate versus inanimate subject referent. If the current interpretation is on the right track, the configuration with an inanimate subject referent would be less costly than the configuration with an animate subject referent in the constitutive-biasing context. Because an inanimate subject referent carries no lexical bias towards a reading that conflicts the contextual constraint, i.e., the lexical representation and contextual constraint both facilitate the constitutive reading, comprehen-

ders should arrive at the constitutive reading easier. In other words, the less conflict between the lexical and contextual constraints (i.e., a conflict of preference), the less processing cost there will be.

The pattern observed here suggests an interplay between lexical representations and contextual constraints, which has been discussed in literature. In particular, the constraint-satisfaction approach claims that competitions between alternative analyses engender computational cost (MacDonald et al., 1994b,a; McRae et al., 1998; Seidenberg et al., 1982; Spivey and Tanenhaus, 1998; Tanenhaus et al., 1979, 1990). The processor takes lexical and contextual constraints into account to find an interpretation that satisfies the constraints. The additional processing effort for AspV sentences results from the competition between alternative interpretations and ambiguity resolution. Besides, different constraints may compete with each other in cases where they favor different representations. This is manifested in the constBiasing-AspV condition where the context favors the constitutive reading yet the lexical bias from an animate subject referent favors the agentive reading. The absence of subordinate effect could result from the interplay between the two kinds of constraints, such that the lexical constraint limits the impact of context.

In summary, the Neutral condition is not harder to process than two Biasing conditions because meaning-frequency—guided by the fact that animate subject referents are associated with controllers more often—prioritizes the

agentive reading in the Neutral condition. Hence, one reading stands out in the Neutral condition, just as the case in the two Biasing conditions. The constBiasing condition is not harder than the Neutral condition and the eventBiasing condition (i.e., no subordinate bias effect) because the biasing contextual information in the former prioritizes the constitutive reading and could attenuate the subordinate bias effect. Taken together, the absence of the context effect is likely to result from the interplay of two factors: (a) biasing contextual information, and (b) meaning-frequency guided by the control asymmetry between the participants in a sentence.

6.5 Conclusion

This chapter examined the context effect on processing AspV sentences with neutral context and two biasing contexts towards the agentive versus constitutive reading respectively in real-time. As expected by the SIH, results show that AspV sentences engendered longer RTs than PsychV sentences.

No effect of context was found, however. I have suggested a few possible explanations for the absence of context effect in ambiguity resolution. It suggests that the dimension composition between an aspectual verb and its complement takes place regardless of context. Also, it is possible that contextual disambiguation may take effect slower than expected and did not accrue in the regions examined here. In addition, the pattern could result from (i)

meaning-dominance of the agentive reading in the AspVanimate configuration, and (ii) the biasing contextual information towards a dis-preferred constitutive reading, given an animate subject referent that is often conceived as a controller. To further tease apart the two factors ((i) and (ii)), it might be useful to test AspV sentences with an inanimate versus animate subject referent following different context types. I leave this investigation to future work.

With this said, one might argue that I failed to find the context effect because the context manipulated in this experiment is not strong enough to influence the interpretation for the AspV sentences. In the next chapter, I will present a questionnaire that examines (i) whether a biasing context guides the ultimate interpretation of AspV sentences, and (ii) whether the context type (biasing vs. neutral) influences the acceptability of AspV sentences. Comparing offline measurements with the current online measurements also helps to clarify the issue regarding the timing of contextual disambiguation. It could be that the context effect takes time to surface, and the context search for disambiguating information is not fully implemented under time pressure. If this is the case, the effect of context modulation should be observed in the offline measurements probed by the questionnaire.

Chapter 7

The context effect on acceptability and the interpretation of sentences with aspectual verbs (Study 4)

7.1 Introduction: Context and sentence interpretation

This chapter tests the second prediction of the Structured Individual Hypothesis (SIH): Biasing contextual information influences the interpretation of sentences with aspectual verbs (AspVs). Recall that the dimension in-

terpretations of AspV sentences with an animate subject referent (i.e., the AspVanimate configuration) can be categorized into two types (Chapter 3, Section 3.2.3 & Chapter 5, Section 5.4.1): (a) The agentive reading along the eventive dimension, in which the subject is perceived as an actor/controller acting on the complement denotation, and (b) the constitutive reading along various dimensions, in which the participants denoted by the subject and the complement are perceived as having symmetric control relation. For instance, sentence (7.1) below can be read as (7.2a) or (7.2b) depending on the context.

(7.1) *William Shakespeare began the volume containing works of classic comedies.*

(7.2) a. → [Agentive reading along the eventive dimension]

William Shakespeare began reading/editing/etc. the volume.

b. → [Constitutive reading along the various dimensions]

William Shakespeare's work was the first subpart of the volume.

If the SIH is right in claiming that the exact interpretation of an AspV sentence is contextually determined, then a prior context biasing towards one particular reading should guide comprehenders towards that reading. In addition, because a disambiguating context aids ambiguity resolution and comprehenders would find it easier to pin down the sentence interpretation, such context is expected to increase the acceptability of the following AspV sentences. This study examines these predictions via a questionnaire that probes

the acceptability and the ultimate interpretations for AspV sentences *in context*.

7.2 Method

7.2.1 Materials

Thirty sets of short passages were created, each consisting of a context sentence and a target sentence.[1] Each target sentence contained either an aspectual verb or a psychological verb; they were adapted from Study 2.[2] The conditions were the same as in Study 3. Each AspV sentence was paired with three types of context sentences, biasing towards the agentive reading along the eventive dimension (**eventBiasing-AspV**) or constitutive reading along informational/spatial dimension (**constBiasing-AspV**), or remaining neutral to allow both readings (**Neutral-AspV**).

In the agentive-biasing context towards the eventive dimension (eventBiasing), an event was mentioned in the context sentence, in which the subject served as an Agent (actor/controller) acting on the entity denoted by the complement. Hypothetically, this would bias the agentive reading along the

[1]The materials used in Study 4 were nearly the same as Study 3 (the eye-tracking experiment); yet the target sentences in Study 3 differ minimally for the purpose of fitting each target sentence in one line on the display screen (13 target sentences were shortened mostly by deleting unimportant adjectives).

[2]The target sentences were modified minimally to bridge the content of the context sentence in this study.

eventive dimension for the following target AspV sentence. In the constitutive biasing context (constBiasing), a structured individual was mentioned (e.g. a collection, a series of X) to facilitate the sub-part reading. For the neutral context, the context sentences were manipulated such that both agentive and constitutive readings were available for the target AspV sentences.[3]

On the other hand, each psychological-verb (PsychV) sentence was paired with two types of context: Neutral (**Neutral-PsychV**) vs. eventive biasing (**eventBiasing-PsychV**). I did not pair PsychV sentences with the constitutive-biasing context because PsychV sentences do not give rise to a constitutive reading. In addition, fifty filler items (same as in Study 3) were introduced. An example set of all conditions is presented in Table 7.1.

7.2.2 Participants

Thirty native speakers of American English were recruited; all between the ages of 18∼30 and without any reading disabilities.

[3]Sentences in the Neutral-AspV condition were checked beforehand by a mini-questionnaire implemented via Yale Qualtrics Survey tools (N=9) to ensure that both agentive and constitutive readings were available for the target AspV sentences. The context sentences were modified afterwards based on the results to maximize equal availability of both readings.

Table 7.1: Study 4—Conditions and sample sentences

Condition	**Context Sentence**	**Target Sentence**
eventBiasing-AspV	In the past, writers often looked at collections of literary work for inspiration.	William Shakespeare began the volume containing works of classic Renaissance comedies.
eventBiasing-PsychV	In the past, writers often looked at collections of literary work for inspiration.	William Shakespeare enjoyed the volume containing works of classic Greek comedies.
Neutral-AspV	Well-known writers usually owned collections of literary work to look at for reference.	William Shakespeare began the volume containing works of classic Renaissance comedies.
Neutral-PsychV	Well-known writers usually owned collections of literary work to look at for reference.	William Shakespeare enjoyed the volume containing works of classic Greek comedies.
constBiasing-AspV	Larry owns many collections of Renaissance literature.	William Shakespeare began the volume containing works of classic Renaissance comedies.
Filler_Ungram	Larry owns many collections of Renaissance literature.	He enjoyed the anthology that goes on a diet.
Filler_Gram	Kevin owns numerous CDs by different jazz musicians.	He loved this CD that his friend gave him.

7.2.3 Procedures

The questionnaire was implemented by Yale Qualtrics Survey Tools. The items (150 critical passages and 50 fillers) were pseudo-randomized for each participant, such that every participant received a unique order of the items

and no consecutive passages were of the same condition. Data collection took place at the Yale Language and Brain Lab.

The stimuli were visually presented on a computer screen. Each trial consisted of (i) passage reading, (ii) acceptability rating, (iii) passage interpretation. Specifically, for each passage consisting of a context sentence and a target sentence, the participants were asked to read the passage, and to rate its acceptability ("*Question: Is the above passage acceptable?*") from a scale 1∼5, defined as below.

5 = I fully understand what you mean and I would express it in that way.

4 = I understand what you mean and I may express it in that way.

3 = I know what you mean, and maybe I could say it in that way.

2 = I am not sure what you mean.

1 = I dont know what you're talking about.

The rating task was followed by a multiple-choice post-hoc question that probed the plausible interpretation(s) of the target sentence in the passage, with following template: [Given the context sentence, what could the target sentence possibly mean?]. Under the question, four types of response options were provided: (a) the agentive reading, (b) the constitutive reading, (c) nonsensical, and (d) others, for which the participants could provide comments in a blank box: [ ]. For each item (i.e., passage), the presentation

order of agentive and constitutive options was automatically randomized by the Qualtrics Survey Tool. The participants were allowed to choose multiple options if appropriate. Below is an example showing the question on sentence interpretation and the options available to the participants.

Given "*Musicians often record their pieces for compilation or memorial albums*", what could the sentence "*Dave Brubeck started this CD of classic Jazz hits*" possibly mean?

◯ *Dave Brubeck started recording this CD of classic Jazz hits.* (Agentive)

◯ *The song by Dave Brubeck was the first track on this CD of classic Jazz hits.* (Constitutive)

◯ *The sentence "Dave Brubeck started this CD of classic Jazz hits" does NOT make sense.* (Nonsensical)

◯ *Others (Please specify in the box):* blank for the participants to fill in

---

7.2.4 Data analysis

Data from twenty-nine participants were analyzed. One participant was excluded because the responses were undifferentiated among the conditions.

I performed mixed-model analyses, implemented in the R statistical environment, using the *lme4* package (Baayen et al., 2008; R Core Team, 2015).

Regarding the rating scores, two sets of analyses were carried out to address two research questions respectively. First, I examined whether the context modulation has any effect on AspV sentences. To this end, I analyzed the rating scores of the three conditions with aspectual verbs: Neutral-AspV, eventBiasing-AspV, and constBiasing-AspV. The context effect on the rating scores was assessed by contrasting a model with context type as the fixed factor against a base model without it. Both models included random intercepts for subjects and items as well as random slopes of context type over subjects and items.[4]

Second, I examined the effects of verb type and context type as well as their interaction. To this end, I analyzed the rating scores of the four conditions as a 2 x 2 design, incorporating two fixed factors, each having two levels: Verb Type (AspV/PsychV) x Context Type (Neutral/eventBiasing). I report results from models that included random intercepts for subjects and items as well as random slopes of context type x verb type interaction over subjects

[4]The maximal models with both random intercepts and random slopes fitted significantly better than the intercept-only models, having smaller AIC values. I therefore reported results from the maximal models.

and items.[5]

Regarding the sentence interpretation task, I calculated the proportion of each response type chosen by the participants per condition. For each AspV condition with a distinct context, I performed a logistic mixed-model analysis to examine the difference between the agentive vs. constitutive reading[6]. A difference in reading type (agentive vs. constitutive) was assessed by contrasting a model with a fixed factor of reading type against a model without it. The models incorporated random intercepts for subjects and items as well as random slopes for reading type over subjects and items.

7.3 Results

7.3.1 Acceptability rating task

The five critical conditions of AspVs and PsychVs were within the acceptable range, all rated above 3.5 in average. The means, standard deviations, standard errors, and medians for each condition are provided in Table 7.2, visualized in Figure 7.1.

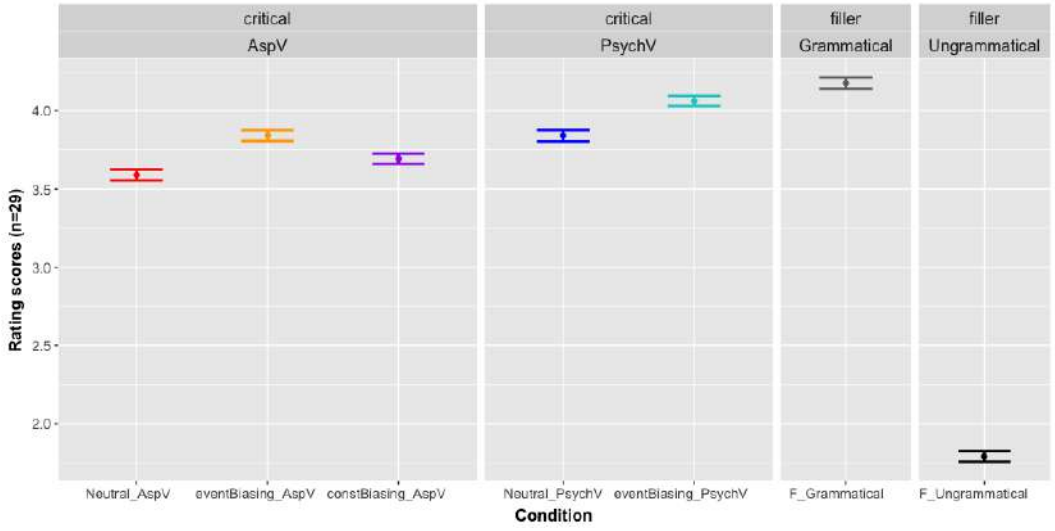
[5]Same as the first set of analyses, the maximal models with both random intercepts and random slopes fitted significantly better than the intercept-only models, having smaller AIC values. The random factors were coded as: `(1 + contexttype*verbtype|Subject) + (1 + contexttype*verbtype|ItemSet)`. Models with random slopes only for subjects but not for items failed to converge. Hence, I report results from the maximal models here; the results remained the same as the intercept-only models.

[6]For this analysis, participants' responses were analyzed as dichotomous, coded as 1 if a reading was selected and 0 otherwise.

Table 7.2: Study 4—Results of the acceptability rating task (n=29)

Condition	**Mean**	**SD**	**se**	**Median**
eventBiasing-AspV	3.84	1.00	0.03	4
Neutral-AspV	3.59	1.03	0.04	4
ConstBiasing-AspV	3.69	0.98	0.03	4
eventBiasing-PsychV	4.06	0.96	0.03	4
Neutral-PsychV	3.84	1.07	0.04	4
Filler_Grammatical	4.18	0.97	0.04	5
Filler_Ungrammatical	1.79	0.91	0.03	2

Figure 7.1: Study 4—Results of the acceptability rating task (n=29). Means with ±1 SE — All conditions



The first set of AspV-only analyses showed an effect of context type ($\chi^2(2) = 16.237$, $p < .001$), in the direction of eventBiasing > constBiasing > Neutral (Figure 7.2, left panel). Pairwise comparisons with Tukey corrections indicated that the eventBiasing-AspV condition was rated significantly higher than the Neutral-AspV condition ($p < .001$); the other two pairwise comparisons (Neutral vs. constBiasing & eventBiasing vs. constBiasing) did not show significance in the maximal model.

The second set of analyses regarding Verb Type (AspV/PsychV) x Context Type (Neutral/eventBiasing) revealed a main effect of Verb Type ($\chi^2(1) = 9.43$, $p = .002$), such that the AspV sentences were rated significantly lower than the PsychV sentences. There was also a main effect of Context Type ($\chi^2(1)=20.877$, $p < .001$), such that the sentences with the Neutral context were rated significantly lower than the sentences with the eventBiasing context (Figure 7.3). No interaction of Verb Type and Context Type was found ($\chi^2(1)=0.14$, $p=0.71$).

7.3.2 Sentence interpretation task

For each condition, I examined the frequency of each response type chosen by the participants; results are presented in Table 7.3. Here I focus on the contextual modulation of the AspV sentences to address the question: Does a prior context sentence biasing towards one particular interpretation have an

Figure 7.2: Study 4—Results of the acceptability rating task (n=29). Means with ±1 SE — The 5 critical conditions (enlarged from Figure 7.1)

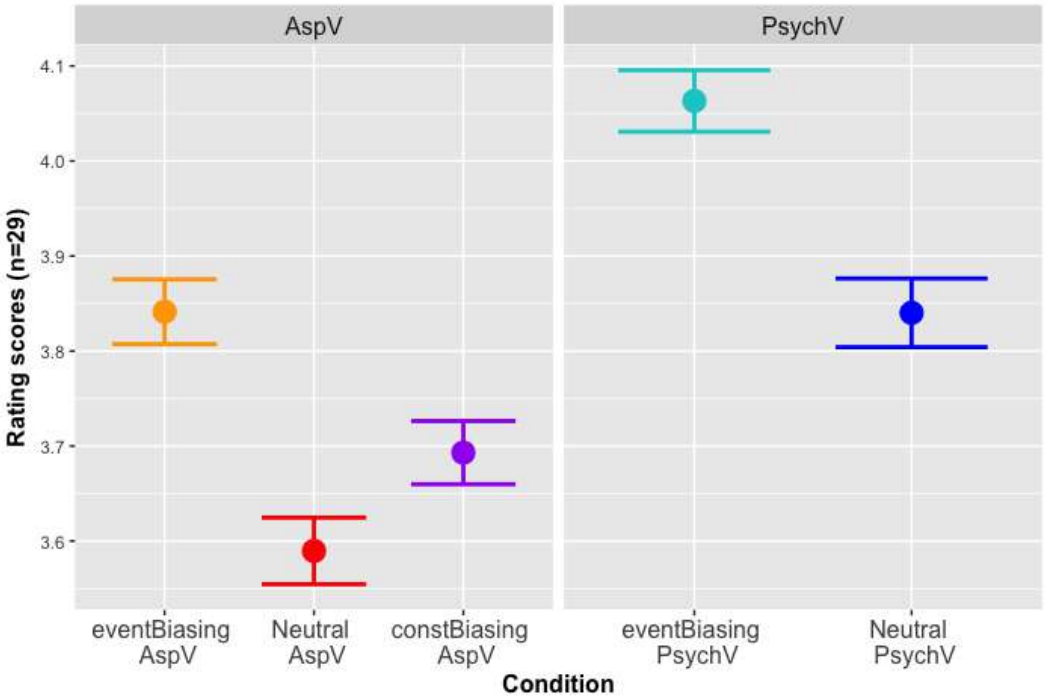
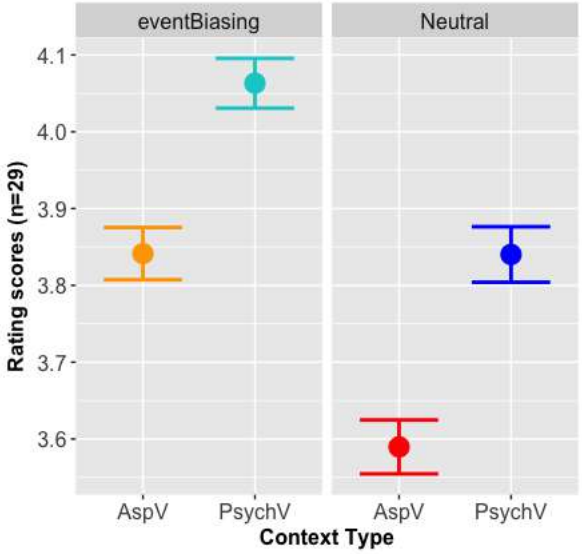


Figure 7.3: Study 4—Results of the acceptability rating task (n=29). Means with ±1 SE — Verb Type (2) x Context Type (2)



impact on comprehenders' interpretation for the AspV sentences?

Table 7.3: Study 4—Percentages of the response types chosen by the participants in the sentence interpretation task

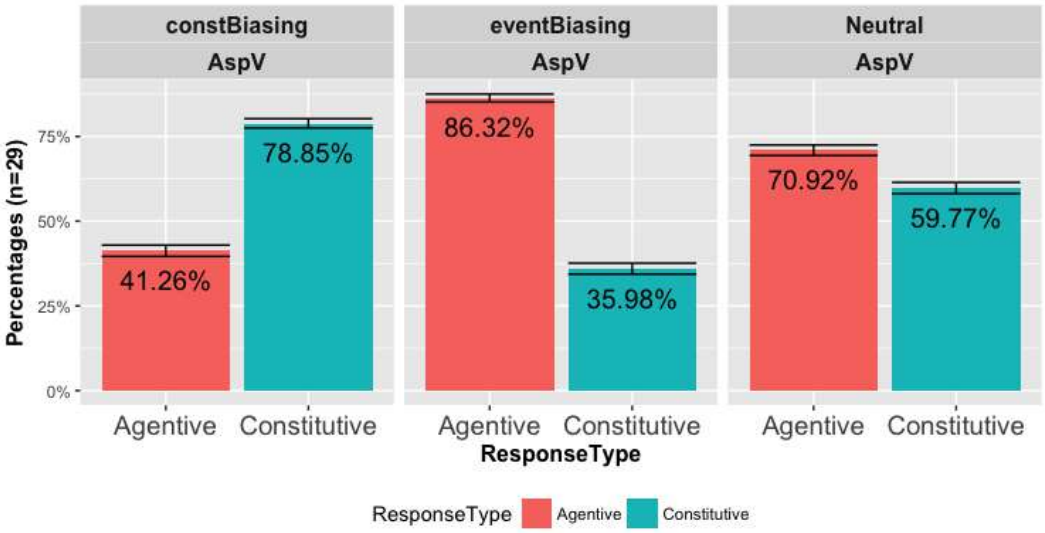
Condition	**Agentive**	**Constitutive**	**Nonsensical**	**Others**
Neutral-AspV	**70.92** %	**59.77**%	3.91%	6.67 %
eventBiasing-AspV	**86.32** %	**35.98** %	2.76 %	4.48%
constBiasing-AspV	**41.26** %	**78.85**%	2.41 %	5.52%
Neutral-PsychV	81.26 %	5.52 %	6.32 %	21.15 %
eventBiasing-PsychV	87.01 %	3.33 %	2.99 %	18.39 %

As shown in Figure 7.4, in neutral context, the target AspV sentences received both the agentive and the constitutive reading with roughly equal percentages, with the [Agentive : Constitutive] ratio being 1.19. For this Neutral-AspV condition, the two readings did not differ significantly ($\chi^2(1)$=1.837, p=0.175).

In the eventive biasing context, the target AspV sentences received the agentive reading twice often than the constitutive reading, with the [Agentive : Constitutive] ratio being 2.40. For this eventBiasing-AspV condition, the two readings differed significantly ($\chi^2(1)$=28.666, $p<$ .001).

In comparison, in the constitutive biasing context, the target AspV sentences obtained the constitutive reading about twice often than the agentive reading, with the [Agentive : Constitutive] ratio being 0.52. For this

Figure 7.4: Study 4—Results of sentence interpretation for the AspV sentences



constBiasing-AspV condition, the two readings differed significantly ($\chi^2(1)$=13.285, $p<$ .001).

On the other hand, PsychV sentences received mostly (over 80%) the agentive reading regardless of context type, for both the Neutral-PsychV and eventBiasing-PsychV conditions. The constitutive (subparthood) reading did not make sense for PsychV sentences. In addition, there were nontrivial proportions for the [Others] response made to both PsychV conditions (1/5 in average). Examining the participants' comments in the blank box provided more closely, I found that the target PsychV sentences often received what I call "the content reading." Some participants indicated that the PsychV sentences could simply refer to the content denoted by the complement itself and not necessarily be related to some event or action. Sample responses noted by the participants included "Dave Brubeck loved the content of this CD of class Jazz hits," for the target sentence "*Dave Brubeck loved this CD of classic Jazz hits*"; "The postdoc may have favored the list of finalists itself," for the target sentence "*The postdoc favored the list of finalist.*" The percentages for the content reading were summarized in Table 7.4, where the percentages represent how often the participants obtained such reading out of all responses made for each PsychV condition.

I turn to our interpretations of these results in Section 7.4 below.

Table 7.4: Study 4—Participants' interpretations for the PsychV sentences

Condition	Content reading	Ambiguous between the eventive and content reading
eventBiasing-PsychV	46.15 %	1.28 %
Neutral-PsychV	35.33 %	0.54 %

7.4 Discussion

Overall, results of acceptability rating show that all five critical conditions were within the acceptable range (> 3.5, while 3 was defined as "*I know what you mean, and maybe I could say it in that way.*") as suggested by the means of rating scores. With respect to the rating scores, sentences with aspectual verbs were rated lower than those with psychological verbs, regardless of context type. This is likely due to the semantic ambiguity involved in AspV sentences, especially in the neutral context where comprehenders can obtain multiple plausible readings and feel uncertain about the exact interpretation. Such uncertainty and the call for resolving semantic ambiguity in AspV sentences could be the source of their lower rating scores.

The effect of context type in the rating task suggests that sentences in the neutral context were rated significantly lower than a biasing context. Importantly, such context effect did not interact with verb type, indicating that a context which provides more information towards a particular reading is more preferable than a context without such information. For AspV sentences, the exact reading is contextually determined according to the SIH. Comprehenders

need to search for relevant information in context to pin down the contextually appropriate interpretation. A richer context increases comprehensibility presumably because it guides comprehenders to arrive at the ultimate interpretation less effortfully by making the relevant information directly available (via linguistic means in the context sentence). In the case of AspV sentences, the relevant information is whether the subject referent is conceived as a controller (taking the Agent role) that acts on the complement denotation in a situation—one that leads to the agentive reading, or whether the two participants are perceived as having symmetric power—a situation that leads to the constitutive reading.

The point here is that, without information that disambiguates sentence meaning, as in the conditions with neutral context, comprehenders need to figure out an appropriate interpretation by themselves. This, I suggest, is done by searching through their conceptual structure to find a coherent mental representation for the sentence in question—a process expected to be costly. The uncertain results of this search process (i.e. multiple plausible interpretations for the sentence) might be the source of the lower rating scores for the neutral context conditions.

On the other hand, the results of sentence interpretation indicate that sentences with aspectual verbs are indeed ambiguous between the agentive reading along the eventive dimension and the constitutive reading along the spatial/informational/temporal/etc. dimensions. Even in the eventive-biasing

context, the constitutive reading was selected for more than 1/3 of the time (35.98%). Likewise, the agentive reading was selected in the constitutive-biasing context with the percentage 41.26%. These strongly support the SIH's claim that the (non-eventive) constitutive reading is available in AspV sentences, contrary to the Type-Shifting account.

Furthermore, the ratios of the two readings for the AspV conditions suggest that the context manipulation was fairly successful. As Table 7.3 shows, the reading type chosen by the participants corresponded to the matching context type. That is, comprehenders obtained the agentive reading along the eventive dimension in the eventive-biasing context twice often than the constitutive reading. In the constitutive-biasing context, comprehenders obtained the constitutive reading mostly.

The findings for the PsychV conditions are noteworthy as well. The results of sentence interpretation indicate that PsychV sentences mainly gave rise to an agentive reading. Although the rating scores showed a context effect for PsychV sentences, such that a biasing context led to higher scores, the sentence interpretation was not modulated by context. This in part due to the fact that PsychV sentences generally do not yield a constitutive reading.s The more prominent finding is the participants' comments that the interpretation of PsychV sentences was not necessarily eventive. For example, "*William Shakespeare enjoyed the volume*" can be properly interpreted as "William Shakespeare enjoyed the *content* of the volume," rather than "enjoyed an activity

associated with the volume." This suggests that psychological verbs, which are previously identified as coercion verbs, do not necessarily require event-denoting complements.

In sum, the results of the sentence interpretation task challenge the assumption of the Type-Shifting Hypothesis, i.e., coercion verbs—including aspectual and psychological verbs—require their complement to denote some event. What is observed here is that sentences with aspectual verbs yielded a non-eventive constitutive reading. More surprising to this account, this reading was available to comprehenders even under the eventBiasing context. Also, sentences with psychological verbs received a non-eventive content reading. These show that sentences with either aspectual verbs or psychological verbs do not always give rise to an eventive reading and hence do not require the type-shifting operation.

On the other hand, the findings are expected by the SIH. Aspectual verbs do not restrict their complements to denote events. Rather, AspV sentences involve semantic ambiguity among multiple dimension readings, which can be categorized into the agentive reading and constitutive readings. The interpretation for an AspV sentence is constrained by context. Comprehenders search for relevant information to pin down the exact dimension along which the sentence is interpreted. If there is no disambiguating information in context, comprehenders could obtain multiple readings so long as they are compatible with context.

7.5 Conclusion

To conclude, Study 4 reveals several points: (1) Sentences with aspectual verbs are indeed semantically ambiguous between an agentive reading along the eventive dimension and a constitutive reading along multiple dimensions. (2) Sentences with psychological verbs do not necessarily give rise to an eventive reading; comprehenders often obtain a (non-eventive) content reading for these sentences. (3) Sentence acceptability is influenced by the richness of the context; passages with relevant contextual information towards one particular reading is preferable than those without it. I suggest that this is because a biasing context reduces the effort of searching through conceptual structure to pin down an exact interpretation. Taken together, the findings support the second prediction of the Structured Individual Hypothesis—disambiguating contextual information guides sentence interpretation.

Chapter 8

Adjudicating between the Type-Shifting Hypothesis and the Structured Individual Hypothesis

8.1 Introduction

In this chapter, I discuss the findings of Study 1∼4 to adjudicate between the Type-Shifting Hypothesis and the Structured Individual Hypothesis. I will argue that the results support the Structured Individual Hypothesis while undermining the Type-Shifting Hypothesis. This suggests that the complement

coercion effect should be viewed as the processing of aspectual verbs (AspVs), which involves meaning underspecification followed by ambiguity resolution, rather than a type-mismatch followed by a repairing operation.

8.2 Adjudicating between the Type-Shifting Hypothesis vs. the Structured Individual Hypothesis

To begin with, let me recapitulate the claims and predictions of the two hypotheses in Table 8.1.

Table 8.1: Type-Shifting Hypothesis vs. Structured Individual Hypothesis

	Type-Shifting Hypothesis	**Structured Individual Hypothesis**
Claims	Coercion verbs, including aspectual verbs (AspVs) and psychological verbs (PsychVs), select for eventive complements, and trigger the type-shifting operation when followed by an entity-denoting complement.	AspVs, but not PsychVs, select for structured individuals, which can be construed along various dimensions, as their complements. AspV sentences are semantically ambiguous among multiple dimension representations, and the ultimate interpretation is contextually determined.
Source of the cost	The repair of the mismatch between a coercion verb and an entity-denoting complement via a type-shifting operation to construct the eventive meaning.	(i) Exhaustive lexical retrieval of an AspV's dimension functions, (ii) Ambiguity resolution among multiple dimension representations via contextual search.
Predictions	Sentences with a coercion verb (AspV or PsychV) + an entity-denoting complement engender additional processing cost.	(i) Sentences with AspVs, but not PsychVs, engender additional cost. (ii) The ultimate interpretation for an AspV sentence is influenced by contextual information.

As pointed out in Chapter 2, the two hypotheses are based on different semantic analyses about the lexical meaning of the verbs. The Type-Shifting analysis identifies a set of coercion verbs—including aspectual verbs and psychological verbs—which select for eventive complements (Pustejovsky, 1991, 1995). The Structured Individual analysis (Piñango and Deo, 2012, 2015) singles out

aspectual verbs, stating that they select for structured individuals as their complements. These two different semantic analyses have different processing implications, leading to the Type-Shifting Hypothesis and the Structured Individual Hypothesis respectively.

These processing hypotheses are tested in Study 1∼4 reported in Chapter 4∼7. The findings of the four studies are summarized in Table 8.2:

Table 8.2: Results summary of Studies 1∼4

	Findings	**Experiment/Task**
(I)	**AspVs > PsychVs:** AspV sentences engendered more processing cost than PsychV sentences—either in isolation or with context.	Self-paced reading, fMRI, Eye-tracking, Acceptability Rating
(II)	**AspVanimate & AspVinanimate pattern alike:** AspV sentences with an animate subject referent (AspVanimate) and those with an inanimate subject referent (AspVinanimate) engendered comparable cost in behavioral measurements, and both induced additional activity in LIFC, as compared to PsychV sentences. In addition, the brain regions recruited by AspVanimate sentences formed a superset of those recruited by AspVinanimate sentences.	Self-paced reading, fMRI
(III)	**Neural correlates of AspVs:** Processing AspV sentences (> PsychV) involved (i) Wernicke's area at Subject∼Verb, and (ii) LIFC at Complement∼S-final	fMRI
(IV)	**Context affects AspV sentence reading (offline):** The interpretation of AspV sentences (agentive or constitutive) was influenced by context. Sentences in a neutral context was less acceptable than those in a context biasing towards either reading.	Sentence interpretation, Acceptability Rating
(V)	**Processing of AspV sentences in context (online):** Following the three types of context (Neutral, eventBiasing, constBiasing), AspV sentences engendered comparable processing cost. This is likely to reflect an interplay of (i) lexical representations of the items in the sentences (ii) contextual constraints in real-time comprehension.	Eye-tracking

Recall the two sets of predictions made by the Structured Individual Hy-

pothesis (SIH) indicated in Chapter 3 (Figure 3.3):

1. **Prediction (I)**: The composition with aspectual verbs leads to greater computational cost, as compared to psychological verbs.

2. **Prediction (II)**: Context constrains the dimension interpretation(s) of AspV sentences and guides ambiguity resolution.

These predictions are supported by the findings, discussed below.

First, the distinct processing profiles—behavioral and neurological—of aspectual verbs versus psychological verbs (Finding I, II, & III) are predicted by the Structured Individual Hypothesis (SIH) but not by the Type-Shifting Hypothesis. The SIH, following Piñango and Deo (2012, 2015), argues that aspectual verbs have to be dissociated from psychological verbs because of their distinct semantic properties. As expected, AspV sentences engendered longer reading times (RTs) in the self-paced reading experiment after the complement head. AspV sentences also engendered longer RTs in the eye-tracking experiment in later measurements at the verb, complement head, and the subsequent region. On the SIH, this is because AspV sentences involve the dimension composition (AspV + complement) and ambiguity resolution among multiple dimension representations. On the Type-Shifting Hypothesis, both AspVs and PsychVs are considered as coercion verbs that trigger the type-shifting operation; this predicts a similar processing pattern for both verb types, contrary to the results.

Second, the similarity of the AspVanimate and the AspVinanimate configurations—Finding (II)—is accounted for by the SIH but is left unexplained by the Type-Shifting Hypothesis. Regardless of subject animacy, AspV sentences in isolation elicited longer RTs in the self-paced reading experiment and additional activity in left inferior frontal cortex (LIFC) after the complement head was encountered, as compared to PsychV sentences. On the SIH, this follows because both configurations of AspVs involve dimension ambiguity (see Chapter 3 & 5). The constitutive reading licensed by both the AspVanimate and AspVinanimate configurations can be interpreted along multiple dimensions (e.g., spatial, informational). Hence, both configurations of AspVs involve ambiguity resolution among multiple dimension representations.

The SIH also explains the neurological finding that the AspVanimate sentences recruited additional brain regions associated with sequential action planning and event-related autobiography than the AspVinanimate sentences (see Chapter 5). The subset relationship in terms of brain activity reflects their relationship in terms of available linguistic interpretations (AspVanimate $\supseteq$ AspVinanimate). As pointed out above, an animate subject referent is a good proxy for a controller and thus leads to the agentive reading in null context, whereas this reading is in general unavailable for AspVinanimate sentences without a supporting context.

On the other hand, the Type-Shifting Hypothesis does not take the AspVinanimate configuration into account at all. The assumption that coercion

verbs (including AspVs) select for eventive complements does not hold in the constitutive reading of AspV sentences (e.g., *"Starry Night" begins the collection of oil paintings.*). Study 2 suggests that the Type-Shifting Hypothesis falls short not only in terms of linguistic coverage but also in terms of processing pattern. The AspVinanimate configuration does not give rise to an eventive interpretation typically, hence no type-shifting operation is needed during comprehension. Yet processing such configuration is as costly as the AspVanimate configuration. If the processing cost associated with AspVanimate sentences—the coercion configuration—results from the type-shifting operation or the construction of an eventive reading, then it is not clear why the AspVinanimate sentences, which do not receive an eventive reading, pattern like the AspVanimate sentences.

Third, AspV sentences induced additional brain activity in Wernicke's area and left inferior frontal cortex—Finding (III). This neurological pattern follows the two hypothesized processes of the SIH: exhaustive retrieval of the lexical functions encoded in AspVs and ambiguity resolution after the dimension composition with the complement. On the Type-Shifting Hypothesis, both AspV and PsychV sentences with entity-denoting complement should call for the type-shifting operation. It is unexpected by the Type-Shifting Hypothesis that the set of coercion verbs exhibited distinct processing profiles. Also, it is unclear how one type-shifting operation is associated with two cortical regions found in the fMRI experiment (Study 1, Exp. 2).

Fourth, the ultimate interpretation of an AspV sentence is modulated by context—Finding (IV). The sentence interpretation task of Study 4 (Chapter 7) demonstrates that AspV sentences with an animate subject referent (the AspVanimate configuration) such as (8.1), are indeed semantically ambiguous between an agentive reading (e.g., 8.2a) and a constitutive reading (e.g., 8.2b) in the neutral context with no disambiguating information.

(8.1) *Dave Brubeck started this CD of classic Jazz hits.*

(8.2) a. Agentive reading: *Dave Brubeck started recording this CD of classic Jazz hits.*

b. Constitutive reading: *The song by Dave Brubeck was the first song of this CD of classic Jazz hits.*

In a biasing context, comprehenders arrive at the reading supported by the context. In Study 4, AspV sentences following a context sentence that biased towards the agentive reading along the eventive dimension (the eventBiasing context) gave rise to the agentive reading. In contrast, sentences following a context that biased towards the constitutive reading along any dimension (the constBiasing context) gave rise to the constitutive reading (see Table 8.3 for example sentences). This is predicted by the SIH: AspV sentences give rise to multiple dimension representations and the ultimate interpretation is contextually determined.

Table 8.3: AspV sentence following different context types (Study 3 & 4)

Context Type	**Context sentence**	**Target AspV sentence**
Neutral	*Many musicians have music libraries that contain tons of albums.*	*Dave Brubeck **started** this CD of classic Jazz hits.*
eventBiasing	*Musicians often record their pieces for compilation or memorial albums.*	
constBiasing	*Kevin owns numerous CDs by different Jazz musicians.*	

This is predicted by the SIH: AspV sentences can give rise to multiple interpretations and the ultimate interpretation is contextually determined.

For AspVanimate sentences, i.e., the coercion configuration (animate subject referent + AspV + entity-denoting complement), the constitutive reading is not only available but even preferred under a supporting context. The Type-Shifting Hypothesis, given its assumption about coercion verbs, does not explain (i) why the coercion configuration turns out compatible with a non-eventive interpretation, and (ii) how different context types induce different interpretations. These findings strongly challenge the Type-Shifting Hypothesis.

Moreover, Study 4 shows that AspVanimate sentences, which license both agentive and constitutive readings, were less acceptable (while still comprehensible) when following a neutral context than when following a biasing context. On the SIH, comprehending AspV sentences involves ambiguity resolution via

contextual search, a process that engenders processing difficulty. A disambiguating context reduces the difficulty of ambiguity resolution by guiding comprehenders towards a particular reading, hence increases sentence acceptability.

Finally, the eye-tracking experiment (Study 3, Chapter 6) shows that AspVanimate sentences elicited comparable RTs (longer than PsychVs sentences) in all context types (eventBiasing, constBiasing, Neutral). This is unexpected by the SIH, which states that dimension ambiguity of AspV sentences is resolved by contextual constraints and hence a biasing context should facilitate the processing of AspV sentences. In Chapter 6 (Section 6.4), I have suggested a few possible explanations for this absence of context. It is likely to reflect the process of dimension composition, which takes place in all cases involving aspectual verbs regardless of context (Chapter 3, Figure 3.2). The impact of context might accrue slower than expected, thus it did not appear in the regions examined in the study. In addition, it suggests that the real-time comprehension of AspV sentences involves an interplay of two components: lexical representations of the items in the sentences and contextual constraints. As argued in Chapter 6, the lack of context effect for AspVanimate sentences could result from the fact that an animate subject referent is more likely to be perceived as a controller (Jackendoff, 2007; MacDonald et al., 1994a; Piñango, to appear). This increases the control asymmetry[1] between the subject and

[1]As a reminder, *control asymmetry* is defined by Piñango (to appear) as "the degree of

the complement denotations, leading to an agentive reading. Therefore, AspVanimate sentences have an inherent preference for the agentive reading in a neutral context, though both agentive and constitutive readings are licensed. I conjecture that AspVanimate sentences in the neutral context patterned like those in the biasing contexts in real-time comprehension due to the preference for the agentive reading in the absence of conflicting contextual constraints (see Chapter 6 for more details). This is likely especially if context constraints take time to accrue.

On the other hand, the Type-Shifting Hypothesis does not explain why AspV sentences in the constBiasing context and those in the eventBiasing context patterned alike. Notice that the AspV sentences in the constBiasing context give rise to a (non-eventive) constitutive reading (evidenced by the sentence interpretation task of Study 4) and hence do not require the type-shifting operation. The Type-Shifting Hypothesis would expect the constBiasing condition to elicit less cost than the eventBiasing condition, which gives rise to an eventive reading via the type-shifting operation, contrary to Finding (V).

Taken together, results obtained from the four studies verify the predictions of the Structured Individual Hypothesis (Table 8.4). The findings are, however, inconsistent with the Type-Shifting Hypothesis.

asymmetric power between potential participants in a relation." A higher control asymmetry refers to a situation in which one participant is capable of exercising power over another participant.

Table 8.4: Comparisons between the Type-Shifting Hypothesis vs. Structured Individual Hypothesis according to the findings

Findings	**Type-Shifting Hypothesis**	**Structured Individual Hypothesis**
(I, II, III, V) AspVs sentences (regardless of subject animacy) require more processing effort than PsychV sentences, either in isolation or with context of different types.		✓ Prediction (I)
(IV) Contextual information guides the interpretation of AspV sentences and influences sentence acceptability.		✓ Prediction (II)

8.3 Conclusion

Given the pattern of the results obtained from the four studies, the phenomenon of complement coercion is better understood as the processing of aspectual verbs, which involves exhaustive lexical retrieval and contextual search as argued by the SIH. The processing of AspV sentences involve meaning underspecification, as the composition of the verb's functions and the complement denotation gives rise to multiple dimension representations. The final interpretation of an AspV sentence is constrained by context.

In the next chapter, I will discuss the implications for the processing models based on the findings, and suggest the components that need to be incorporated in a processing model for meaning underspecification.

Chapter 9

Implications for processing models of meaning underspecification

9.1 Introduction

In this chapter, I discuss the implications for models of sentence processing as suggested by the Structured Individual Hypothesis (SIH). In Chapter 8, I show that the findings support the SIH and that the complement coercion effect should be viewed as the processing of AspV sentences. As AspV sentences involve meaning underspecification, their processing profile informs how comprehenders obtain underspecified meaning in real-time. In what follows, I

suggest the properties of processing AspV sentences (Section 9.2). These properties should be taken into consideration by any model for processing meaning underspecification. Furthermore, the processing architecture implicated by the SIH is consistent with the constraint-satisfaction approach to real-time sentence comprehension (Section 9.3).

9.2 The processing system for meaning underspecification as suggested by the Structured Individual Hypothesis

I hold that the goal of sentence comprehension is to establish a mental representation that fits the context in which the sentence appears. According to the SIH, the comprehension of meaning underspecification involves two processes: (a) exhaustive retrieval of lexical representations, and (b) ambiguity resolution based on contextual constraints. This means that sentence processing involves lexical and contextual components. I discuss these components in Section 9.2.1 & Section 9.2.2 respectively, using AspV sentences to illustrate. In Section 9.2.3, I suggest how these components are incorporated in a processing model of meaning underspecification.

9.2.1 Component 1: Lexical representations

The component of lexical representations in comprehension can be teased apart into a few aspects: (i) satisfaction of lexical requirements, (ii) exhaustive lexical retrieval, and (iii) activation degrees of different representations dependent on meaning frequency.

First, the requirements of lexical representations must be fulfilled during sentence comprehension in order to yield a sensible interpretation. For example, aspectual verbs require their complement to be conceived as a structured individual, which is construed as a directed path structure along some dimension. This conceptualization is the foundation of the meaning of AspV sentences—a subpart relation between two entities (Piñango and Deo, 2012, 2015). Requested by an aspectual verb, the processor conceptualizes the verb's complement as a structured individual; the lexical constraint is thereby fulfilled during the composition with the complement.

Second, when a lexical item is received during comprehension, all lexical information associated with it will be retrieved. The SIH states that the processor exhaustively retrieves all lexical functions (e.g., f_{space}, f_{time}, f_{event}, f_{info}) encoded in AspVs. In Study 1, the greater activation of the supramarginal gyrus (BA 40) for AspV sentences (> PsychV sentences) during the Subject+Verb composition likely reflects the retrieval of multiple lexical functions encoded in aspectual verbs. This is consistent with several behavioral

and neurological studies, which have suggested that the processing demand of verbs increases with more complementation options they encode (e.g., Shapiro et al., 1989; Shetreet et al., 2007, 2016; Swinney, 1979; Thompson et al., 1997).

In particular, Shetreet et al. (2016) show that the left STG activity is greater for verbs with multiple complementation options—defined either syntactically as subcategorization frames or semantically as thematic frames—than verbs with one option. For example, the verb "*remember*" takes a NP/theme as its object or an embedded clause/proposition (9.1a-b), while "*punish*" takes an NP/patient as its complement (9.2).

(9.1) a. *Dan remembered* [*Jenny*].

b. *Dan remembered* [*that Jenny is coming tonight.*]

(9.2) *Dan punished Jenny.*

—Examples from Shetreet et al. (2016)

They argue that verb complementation options can be viewed as a lexical-syntactic ambiguity with respect to the complement type. Exhaustive access applies to the lexical-syntactic ambiguity of a verb's complementation options, just as it applies to the lexical-semantic ambiguity of a noun. In either case, the more lexical structures an item has, the more retrieving effort it requires.

While lexical retrieval is exhaustive, the degree of activation for each lexical meaning may vary depending on frequency, as suggested by the constraint-satisfaction approach (e.g., Jackendoff, 2007; MacDonald et al., 1994b; Pearl-

mutter et al., 1994; Seidenberg et al., 1982). For instance, an animate subject referent takes the Agent role more often than the Theme role in neutral context, hence the thematic structure with an agentive subject will be activated to a greater degree or faster than the structure in which the subject takes the Theme role. Suppose the processor evaluates the alternative semantic representations in the order of their activation levels (Braze, 1999; Pearlmutter et al., 1994; Seidenberg et al., 1982; Tanenhaus et al., 1979); the more strongly activated (i.e., the more frequent) representation will be available to the processor more rapidly. If this representation satisfies the contextual constraints (consistent with the contextual information), then it will be adopted by the processor. In contrast, if the most activated representation is inconsistent with the context, the processor will adopt the next available alternative (which is less activated but licensed by the lexical representations) which satisfies the contextual constraints. In the latter case, lexical meaning frequency and contextual constraints favor different representations. It manifests what Seidenberg et al. (1982, p.532) refers to as "a garden path where the most probable reading is contextually inappropriate." This shows an interplay between lexical representations and contextual constraints.

The above discussion suggests a processing system that involves *exhaustive lexical retrieval, with the activation levels for distinct lexical meanings dependent on frequency.* These processing properties provide an explanation for why AspVanimate sentences (with an animate subject referent) in the neu-

tral context pattern like the those in a biasing context in Study 3 (Chapter 6, Section 6.4.1). Consider the following. (i) An animate subject referent is associated with a controller more frequently, hence carries a bias towards a higher control asymmetry between the subject and the complement denotations, leading to the agentive reading. (ii) The contextual information in the neutral context is compatible with the agentive reading. Suppose that for AspVanimate sentences, the agentive reading is activated to a greater degree than the constitutive reading due to the meaning frequency associated with the animate subject, and suppose that the processor evaluates the alternative semantic representations in the order of activation degree. In processing AspVanimate sentences in the neutral context, the processor will take the agentive reading since it is evaluated first and consistent with the context. Therefore, in both the neutral context (Neutral) and the context that biases towards the agentive reading (eventBiasing), the agentive reading is favored, resulting in a similar pattern of the two conditions. This suggests that in the absence of a conflicting contextual constraint, the lexical preference guides the reading.

Furthermore, the lexical representation of a complement that follows an aspectual verb may constrain the dimension readings of an AspV sentence. For instance, the denotation of "*the banquet*" is usually mapped along the eventive dimension and not the spatial dimension. This explains why "*started the fight*," which has a strong bias towards the eventive dimension, is easier to process than "*started the puzzle*," which is ambiguous among even-

tive/spatial/information dimensions, reported in Traxler et al.'s (2002) study. I conjecture that when a complement's lexical representation licenses multiple dimensions, they may vary in the degree of activation dependent on the frequency of usage. For example, the denotation of "*the party*" is mapped along the eventive dimension *preferentially* though it can also be mapped onto the temporal dimension; in comparison, "*the high way*" is mapped onto the spatial dimension preferentially. Notice that the dimensions are not mutually exclusive to each other necessarily. For a sentence like (9.3), the complement (*the book*) can be interpreted along the informational dimension (9.4a) and the spatial dimension (9.4b) simultaneously.

(9.3) *The author's bibliography begins the book.*

(9.4) a. Informational dimension: the book refers to a body of informational content.

(Reading: The author's bibliography is the initial subpart of the content of the book.)

b. Spatial dimension: the book that refers to a physical entity.

(Reading: The author's bibliography is printed on the first page of the book.)

Taken together, the processing of AspV sentences suggests exhaustive lexical retrieval, with the degree of activation dependent on the frequencies of the alternative dimension representations. For a sentence to yield a compre-

hensible interpretation, the lexical semantic requirements must be fulfilled. In the absence of contextual constraints, semantic ambiguity is guided by lexical representations.

9.2.2 Component 2: Contextual constraints

The comprehension of meaning underspecification involves not only exhaustive lexical retrieval but also ambiguity resolution based on context when the composition of lexical representations gives rise to multiple semantic representations. On the SIH, the contextual constraints for comprehending AspV sentences involve: (a) the control relation between the participants denoted by the subject and the complement, and (b) the choice of dimension (Chapter 5, Table 5.1).

For AspV sentences, the control relation between the subject and complement denotations perceived in a particular context disambiguates the agentive vs. constitutive reading. If the control relation between the two participants is perceived as asymmetry, an AspV sentence gives rise to an agentive reading, in which the subject referent is conceived as a controller performing some activity associated the complement. If the control relation is perceived as symmetric, an AspV sentence gives rise to a constitutive reading, in which the subject denotation is mapped onto a subpart of the structured individual denoted by the complement.

Importantly, the control relation does not hinge on subject animacy. An AspV sentence with an animate subject referent can engender a constitutive reading, as in "*The tall man began the queue in front of the famous bakery.*" Reversely, an inanimate subject referent can yield an agentive reading, as in "*The printer started my documents two minutes ago.*" I have argued that the control relation perceived in context predicts sentence reading better than subject animacy does, because the mapping between the asymmetric vs. symmetric control relation and the agentive vs. constitutive reading is one-to-one (Chapter 5, Section 5.4.1). Therefore, it is more beneficial for the processor to pin down the control relation between the participants to obtain the exact sentence reading. I suggest that the processor figures out the control relation by mining lexical representations and relevant contextual information. The lexical representation of an animate subject could give the processor an initial preference for a higher control asymmetry, but this preference can be overwritten later on by other lexical information, such as the complement's representation, and/or contextual information available to the processor.

Here is an illustration. The sentence "*Dave Brubeck began this CD of classic Jazz hits.*" is ambiguous between an agentive reading (e.g. *Dave Brubeck began recording this CD...*) and a constitutive reading (e.g. *Dave Brubeck's song was the first song on this CD...*). When the processor encounters the subject, its lexical representation would bias the processor towards a control asymmetry and hence the agentive reading, which is supported by a neutral

context or a context biasing towards that reading. However, if the sentence is processed in a context supporting the constitutive reading (e.g. *Kevin owns numerous CDs by different Jazz musicians.*), then the control relation between the participants (*Dave Brubeck & this CD*) is perceived as symmetric, yielding the constitutive reading. In this case, the contextual information that supports the constitutive reading surpasses the initial preference of the subject denotation which favors the agentive reading. This shows that contextual constraints guide the processor towards the contextually supported reading.

Another component that the processor pins down in context is the *dimension* along which the structured individual denoted by the complement is construed. For example, in "*The acknowledgement ends the book*," both the spatial dimension and the informational dimension are plausible. For the dimension ambiguity, the processor needs to determine the dimension that is most relevant in a given context. For instance, the spatial dimension, along which the complement refers to a physical entity, is chosen if the sentence appears in a conversation between two printers while working:

(9.5) A: "*I must do it wrong; the author's bibliography was accidentally printed on the final page of the book.*"

B: "*No! For the books by Publisher X, the acknowledgement ends the book, not the bibliography. We need to reprint it again.*"

If the sentence appears in a conversation between two authors while discussing

how the content of the book should be organized as in (9.6), the informational dimension will be chosen, along which the complement refers to a body of informational content.

(9.6) A: "*Should we put the authors' notes at the end of the book?*"

B: "*Hmm.... According to my understanding of Publisher X, usually the acknowledgement ends the book.*"

The above examples demonstrate that while the composition of the verb and the complement gives rise to multiple dimension representations, context will constrain the dimension along which the sentence is interpreted.

In short, the contextual constraints that determine the exact interpretation of AspV sentences include (i) the control relation between the subject and complement denotations as perceived in a particular context, and (ii) the specific dimension relevant in context. The search for contextual information can be viewed as a process of contextualization—finding a mental representation coherent with the context, given the linguistic input. This contextual search is required to resolve semantic ambiguity and incurs additional processing cost as shown in the four studies.

9.2.3 Properties of the processing system for meaning underspecification

In the above subsections, I have pointed out the lexical and contextual components involved in the processing of AspV sentences. These components are taken into consideration by the processor to deal with underspecification among multiple semantic representations of a sentence. Based on the observations from AspV sentences, I suggest that the processing of meaning underspecification involves the following properties:

(a) Lexical representations, such as a verb's lexical functions and the dimensions licensed by a complement's denotation, are exhaustively retrieved. Lexical semantic requirements, such as AspVs' requirement for a structured individual, need to be satisfied within the sentence (in this case, by conceiving the complement denotation as a structured individual).

(b) The composition of lexical representations gives rise multiple semantic representations *in parallel*. The processor does not immediately commit to one semantic representation; instead, it leaves the semantic representations underspecified, remaining open to all possibilities during the composition.

(c) Multiple semantic representations engendered by the composition are

narrowed down by contextual constraints. Non-supported representations are pruned as comprehension unfolds.

(d) Lexical representations, such as the frequencies of different control relations (e.g., an animate subject referent increases the possibility of a control asymmetry that leads to an agentive reading), may limit the effect of contextual constraints in cases where the two kinds of constraints favor different representations.

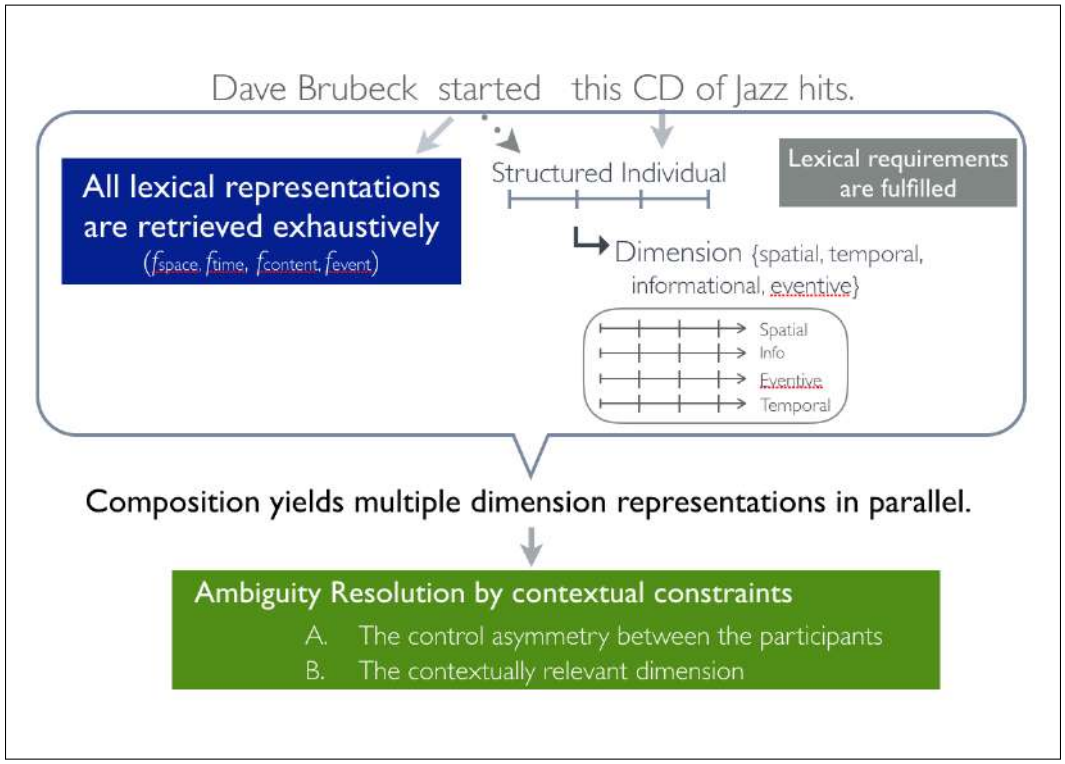
Points (a) and (b) are hypothesized by the SIH and evidenced by the finding that AspVs induced additional cost than PsychVs (Study 1 & 2). Point (c) is also contended by the SIH and evidenced by the finding that comprehenders arrived at the interpretation supported by context (Study 4). Point (d) is suggested by Study 3, motivated to capture the absence of context effect in real-time comprehension (i.e., the eventBiasing, constBiasing, Neutral conditions of AspVs patterned alike) while the verb type effect (AspV > PsychV) was found consistently.

Here is an example that demonstrates these properties. When the processor encounters an aspectual verb (e.g., *start*), it retrieves all of its lexical functions (e.g., f_{space}, f_{time}, f_{info}, f_{event}). As the AspV requires its complement to denote a structured individual, the processor imposes this requirement to the following complement (e.g., *this CD*), conceptualizing it as a structured individual. When the complement is encountered, all dimensions available in

its lexical representation are retrieved (e.g, informational, spatial, eventive, temporal). The structured individual denoted by the complement is mapped onto the dimensions by the verb's lexical functions, construing a directed path structure along various dimensions. In the meantime, the verb (*start*) maps the the subject denotation onto the initial subpart of the directed path structure construed by the complement (*this CD*). The composition of the verb and the complement gives rise to multiple dimension representations (e.g., *this CD* refers to a physical entity or a body of informational content). Facing this ambiguity among multiple dimension representations, the processor examines the contextual constraints, including the control relation between the participants and the specific dimension relevant in the current context. The representations that is not supported by the context are pruned, while those compatible with the context are retained. The processor arrives at an interpretation that satisfies the lexical and contextual constraints. Figure 9.1 illustrates these processes with an example.

In this picture, contextual constraints surface later than lexical constraints, as the composition of lexical semantic representations is assumed to take place regardless of context. The effect of context could be affected by the frequency of alternative lexical meanings in real-time comprehension. As said, an animate subject referent is perceived as a controller more often and hence the agentive representation will be activated to a greater degree or become available to the processor faster. The lexical bias towards the agentive reading

Figure 9.1: Real-time comprehension of meaning underspecification, exemplified by an AspV sentence



could limit the context effect when the context biases towards the constitutive readings. The competition between the constraints which prioritize different alternatives is expected to engender additional processing cost, as discussed in Chapter 6 (Study 3).

As Figure 9.1 shows, the processing model suggested by the SIH is rooted in lexical semantic representations and constrained by context. Processing cost results from exhaustive lexical retrieval and ambiguity resolution among multiple semantic representations. In this view, sentence comprehension involves no mismatch situation (i.e., no misalignment of lexical semantic representations), hence there is no need to introduce any repairing operation (contrary to the Type-Shifting Hypothesis). I suggest that the properties identified above (a-d) should be taken into consideration by any processing model that aims to capture meaning underspecification, given that they are shown by the processing of AspV sentences. Although this view of processing underspecified meaning is motivated by the comprehension of aspectual verbs, I think it is generalizable to other cases (e.g., aspectual coercion) with some modifications.

9.3 The Structured Individual Hypothesis and the constraint-satisfaction approach

The processing view implicated by the SIH is consistent with the constraint-satisfaction approach—particularly the model proposed by Crain and Steedman (1985) for local syntactic ambiguity (i.e., garden path sentences that are temporarily ambiguous between the complement and relative clause readings, as in "*The horse raced past the barn fell.*"). In Crain & Steedman's weakly interactive model, the syntactic processor constructs several alternative structures in parallel, which are evaluated by reference to context and plausibility, defined as "the likelihood of a particular reading in the light of general real-world knowledge."[1] The model is "weakly" interactive because semantics and context only dispose among the alternative proposed by the syntactic processor, rather than determining *which* structures are proposed in the first place (as in strongly interactive models).

According to Crain and Steedman (1985), the phenomena taken to reflect parsing strategies such as Minimal Attachment in the serial models (Frazier

[1]Specifically, they indicate that "local syntactic ambiguities are resolved by an interaction of the 'weak' variety from semantics and reference, which can evaluate alternatives proposed by the syntactic processor" (Crain and Steedman, 1985, p.337).

and Fodor, 1978; Frazier, 1979) in fact result from the evaluation of alternative structures based on the relative goodness of fit to the context (p.345). They show that the garden path effect (the additional processing cost caused by local syntactic ambiguities) can be attenuated or created by context manipulation, such as the availability of biasing information, plausibility of a particular reading with respect to general world knowledge, and definiteness of a noun phrase (NP) that carries referential, presuppositional characters. Consider their examples in (9.7) and (9.8); sentences (9.7) are contrasted with (9.8) in definiteness, while sentences in (a) are contrasted with those in (b) in plausibility.

(9.7) *(a) The teachers/(b) The children taught by the Berlitz method passed the test.*

(9.8) *(a) Teachers/ (b) Children taught by the Berlitz method passed the test.*

—Examples from Crain and Steedman (1985, p.339)

Using a grammatical judgement task, they found that the more plausible sentences (b) were judged grammatical significantly more often than the less plausible sentences (a). Meanwhile, the indefinite sentences (9.8) were judged grammatical significantly more than the definite sentences (9,7), because indefinite sentences require fewer presuppositions of the subject NPs.

In another experiment, they show that grammatical judgement for target sentences which are temporarily ambiguous between the complement and

relative-clause readings (e.g., the underlined part of "*The psychologist told the wife that he was having trouble with* [*her husband/to leave her husband*]*.*") was influenced by contexts that biased distinct readings. The inconsistent context-target pairings elicited more ungrammatical responses. These findings suggest that contextual information is used to compare between multiple representations. The processor chooses the representation that has the fewest unsatisfied constraints[2].

The processing view implicated by the SIH for meaning underspecification is in line with this constraint-satisfaction model. On the SIH, the processor constructs multiple semantic representations in parallel based on the received lexical representations, and evaluates the alternative representations based on context (Figure 9.1). While Crain and Steedman (1985) suggest that all syntactic structures are constructed simultaneously and evaluated by semantics and context, the SIH suggests that all semantic representations are constructed simultaneously and evaluated by contextual constraints. In other words, the processing model implicated by the SIH parallels Crain & Steedman's model; the two differ in the phenomena they target—semantic underspecification and local syntactic ambiguity respectively. Below I list several similarities between the processing view of the SIH and that of the constraint-satisfaction approach more specifically.

[2]Crain and Steedman (1985) argue that there is no intrinsically garden-path structures which cannot be eliminated by semantics and context. Rather, sentences show the garden-path phenomenon in certain contexts but not others.

First, the SIH and the constraint-satisfaction approach hold that the processor constructs multiple representations *in parallel* rather than building one representation according to certain domain-specific strategies (e.g., Minimal Attachment, Late Closure as in the serial models). The processor does not commit to single representation immediately, leaving the representations underspecified before they are disambiguated (when one representation wins out).

Second, both the SIH and the constraint-satisfaction approach highlight the role of context, which is used to evaluate alternative representations in order to pin down an appropriate interpretation for the sentence in question. Crain and Steedman (1985) suggest that semantic and contextual information (e.g., definiteness of a noun phrase, plausibility of a reading in a particular context) is used to accept or reject alternative analyses (the relative-clause analysis vs. complement analysis) for garden-path sentences on a word-by-word basis. Similarly, the SIH states that the processor considers contextual constraints (e.g., control relation, the relevant dimension) to determine the ultimate semantic representation. In both cases, the processor chooses the one that has the least conflict with the context. The mechanism of constraint checking based on context applies to the processing of meaning underspecification in addition to the processing of syntactic ambiguity.

Third, both attribute additional processing cost to choosing one representation among competing alternatives as a process of ambiguity resolution. On the constraint-satisfaction approach to the processing of syntactic ambiguity,

difficulty arises when the processor needs to determine one analysis among multiple syntactic structures. In parallel, the SIH for meaning underspecification holds that the determination among multiple semantic representations engenders cost.

Fourth, both the SIH and the constraint-satisfaction approach consider ambiguity as rooted in lexical representations. The constraint-satisfaction approach holds that the existence of multiple argument structures stored in lexical entries results in lexical ambiguities. On the SIH, the composition of the verb's multiple lexical functions and the complement gives rise to multiple representations in the case of AspV sentences. The agentive and the constitutive readings of these sentences correspond to two different thematic structures as shown in Table 9.1. This instantiates a lexical ambiguity according to the constraint-satisfaction approach. That is, an ambiguity arises because lexical items license multiple representations, hence it is lexical in nature.

Table 9.1: Thematic structures of AspV sentences in different readings

Reading	Thematic structure <Subject, Object>	Example
Agentive	<Agent, Theme>	*The princess began (reading) the book.*
Constitutive	<Theme, Location>	*The princess('s story) began the book.*

In addition, the lexical properties suggested by the constraint-satisfaction approach explain the data of Study 3 (eye-tracking). I have mentioned in Section 9.2.1 that the degree of activation for distinct lexical meanings may be frequency-dependent, as suggested by constraint-satisfaction models (e.g.,

Jackendoff, 2007; MacDonald et al., 1994a; Pearlmutter et al., 1994; Seidenberg et al., 1982). For instance, the frequencies associated with different argument structures of a verb could bias the processor in selecting between alternative representations. Study 3 shows no difference between the event-Biasing vs. Neutral context for AspVanimate sentences. According to the constraint-satisfaction approach, this results from the unbalanced frequencies of the different argument structures, given that an animate subject referent is more likely to take the Agent role than the Theme in the absence of conflicting contextual information. By hypothesis, the two alternative argument structures are activated with different degrees according to their frequencies of use, though both are retrieved when the verb is encountered. Because both the eventBiasing and Neutral contexts support the agentive reading (no violation of contextual constraints), and because this reading is favored by meaning frequency of the lexical representations, the processor adopts the agentive reading in both conditions. As a result, no difference between the two was found.

Finally, the constraint-satisfaction approach emphasizes that constraints may interact and limit each other's impact during comprehension. This accounts for the pattern observed in Study 3—this time regarding the no difference between the Neutral-AspV and constBiasing-AspV conditions. The absence of context effect could result from the interplay of an animate subject referent and the contextual constraints, as the two favor different representations. (*cf.* MacDonald et al., 1994b). The animate subject referent of the

target AspV sentences favors the agentive reading (meaning dominance of the lexical representation), while the constBiasing context favors the constitutive reading (the contextual constraint). As MacDonald et al. (1994b, p.139) put: "Thus, properties within the lexicon, here, sensitivity to frequency, limit the effects of even strong biasing context." The processing profile of AspV sentences in different contexts reflects an interaction between lexical representations and contextual constraints.

In short, the SIH aligns with the constraint-satisfaction approach in (i) allowing the processor to construct multiple representations in parallel without committing to one immediately, (ii) highlighting the role of context in evaluating alternative representations during comprehension, (iii) attributing additional computational cost to the determination of an appropriate representation among multiple alternatives, (iv) treating ambiguity as rooted in lexical representations, and (v) assuming that lexical and contextual constraints work in tandem, interacting or competing with each other during processing. The SIH's view for the processing of meaning underspecification parallels the constraint-satisfaction approach to the processing of (local) syntactic ambiguity.

9.4 Conclusion

In this chapter, I have identified the components involved in the processing of meaning underspecification. The properties revealed by the comprehension of AspV sentences suggest a processing model that is lexically rooted and contextually constrained (*cf.* MacDonald et al., 1994b, p.140). During comprehension, the processor exhaustively retrieves lexical representations. The lexical requirements imposed by the received items must be satisfied within the sentence. The composition of lexical representations gives rise to multiple semantic representations simultaneously. The alternative representations are constrained by context, such that the unsupported representations are pruned. That is, semantic ambiguity is resolved as a process of constraint satisfaction. I suggest that context provides the processor relevant information (e.g. control relation & the specific dimension) to evaluate, rather than pre-selecting, the alternative semantic representations. In addition, lexical and contextual constraints may limit each other's impact in real-time processing in cases where they favor different alternative representations. The goal of the processor is to find a mental representation that satisfies lexical constraints and contextual constraints maximally for a given sentence.

This processing view, suggested by the SIH, is in contrast to the *mismatch*→ *repair* view—one that is implied by the Type-Shifting Hypothesis (Chapter

1). I argue that the processing of underspecified meaning involves no misalignment of lexical semantic requirements (type mismatch), hence repairing operations are not necessary. The computational cost associated with semantic underspecification results from ambiguity resolution among multiple semantic representations[3] (instead of repair), alongside exhaustive lexical retrieval and the composition of lexical representations.

[3]Notice that the semantic ambiguity identified by the SIH is between multiple dimension representations (e.g., the agentive along the eventive dimension vs. constitutive reading along spatial/informational dimensions of AspV sentences). This is different from the ambiguity regarding the activities associated with the complement denotation (e.g. *reading/writing/editing the book*). As Frisson and McElree's (2008) study shows, the cost associated with complement coercion is not modulated by the preference of a specific activity associated with the complement. On the SIH, the possible activities are underspecified in the agentive reading along the eventive dimension.

Chapter 10

Conclusion

10.1 Summary & Contributions

In this dissertation, I examined the complement coercion phenomenon, in which an entity-denoting complement receives an eventive reading when it follows verbs such as *enjoy* and *begin*, as in "*The professor began/enjoyed the volume.*" Sentences involving complement coercion contain a piece of meaning that is not morpho-syntactically supported and require additional effort in real-time comprehension. Investigating the source of the associated cost informs how comprehenders obtain unstated meaning in real-time.

I have compared two hypotheses that treat the complement coercion phenomenon in different ways. The Type-Shifting Hypothesis assumes that coercion verbs such as *enjoy* and *begin* select for an eventive complement. Com-

bining these verbs with an entity-denoting complement results in a semantic mismatch that calls for a repairing type-shifting operation. This hypothesis contends that the additional cost associated with complement coercion results from the insertion of this operation for generating an eventive reading during comprehension . The Type-Shifting Hypothesis implicates a processing system that involves a *mismatch-repair* mechanism for comprehending unstated meaning. It suggests a *semantic garden path* view that attributes processing cost to a misalignment of lexical semantic requirements followed by a repair.

On the other hand, the Structured Individual Hypothesis (SIH), following (Piñango and Deo, 2012, 2015), treats the complement coercion phenomenon as the processing of aspectual verbs (AspVs, e.g., *begin, finish*). Sentences with aspectual verbs involve meaning underspecification, giving rise to a semantic ambiguity among multiple dimension representations—including an agentive reading along the eventive dimension (e.g., *The princess began reading/writing/etc. the novel.*) and a non-eventive, constitutive reading along various dimensions (e.g., *The princess's story/utterance/etc. began the novel.*). On this hypothesis, the additional processing cost that is previously argued to reflect complement coercion in fact results from (i) exhaustive retrieval of the verb's lexical functions that map the complement onto various dimensions, and (ii) ambiguity resolution among multiple readings based on context. These processes are associated with aspectual verbs only, hence these verbs are dissociated from other verb types according to the SIH. The SIH implicates a

processing system that involves no semantic mismatch, therefore no repairing mechanism is required (in contrary to the Type-Shifting Hypothesis). It suggests an underspecification and constraint-satisfaction view that the processor constructs multiple semantic representations in parallel, which are further constrained by context. The determination of a contextually appropriate representation among multiple alternatives engenders processing cost.

I reported four studies to adjudicate between the Type-Shifting Hypothesis and the Structured Individual Hypothesis (Chapter 4∼7). Results show that (i) AspV sentences are more costly than psychological-verb sentences, (ii) processing AspV sentences preferentially recruits Wernicke's area and left inferior frontal cortex, argued to reflect lexical retrieval and ambiguity resolution respectively. (iii) AspV sentences are indeed semantically ambiguous and the ultimate interpretation is influenced by context, and (iv) the impact of context could be limited by lexical representations in real-time comprehension. I argue that the findings support the Structured Individual Hypothesis. The complement coercion phenomenon should be viewed as the processing of aspectual verbs, which involves no semantic mismatch-repair but meaning underspecification and ambiguity resolution.

These findings shed light on the mechanisms implemented in human mind for sentence comprehension. The processing system for meaning underspecification exhibits the following properties: the processor exhaustively retrieves all lexical information encoded in each lexical item for constructing a mental

representation for the sentence in question. The specific lexical requirements (e.g., AspVs select for a structured individual as their complement) must be fulfilled. The composition of lexical representations yields multiple semantic representations in parallel, and the processor searches for relevant contextual information (e.g., control relation between the participants and the dimension referred to by the context) to determine a contextually appropriate representation for the sentence. This is consistent with the constraint-satisfaction approach in that (i) no repair or reanalysis is involved in comprehension, and (ii) additional processing cost results from the competition among alternative representations (Crain and Steedman, 1985; MacDonald et al., 1994a,b; Pearlmutter et al., 1994; Seidenberg et al., 1982; Trueswell and Tanenhaus, 1994).

In sum, the processing system revealed by this dissertation work is essentially rooted in lexical representations. There is no need to invoke additional interpretive operations extraneous to lexical representations (i.e., operations that manipulate lexical information as input), because the information needed for obtaining sentence meaning can be found in lexical representations. In semantic processing, the processor forms multiple semantic representations without committing to a single representation immediately, and seeks to pin down an appropriate interpretation that fulfills all lexical requirements and satisfies contextual constraints maximally.

10.2 Future directions

There are several issues that I have not been able to address or clarify in this dissertation due to time constraints. I list them below to provide future directions.

First, the context effect on AspV sentences with animate subject referents (AspVanimate) that is observed in the offline measurements (Study 4, questionnaire) but not in the online measurements (Study 3, eye-tracking) needs further investigation. Although I have provided a few possible explanations, they have to be verified by future studies. Specifically, I suggested in Chapter 6 that the dimension composition of AspV sentences occurs regardless of context and the impact of contextual constraints takes time to emerge[1]. Therefore, the context effect appeared in the offline tasks (acceptability rating & sentence interpretation) but not in the online reading task of the eye-tracking experiment. It could be that the sentences used in the eye-tracking experiment (Study 3) are too short for the effect to accrue at the regions examined. This can be tested by a modified experiment with longer target sentences. If the interpretation is on the right track, we should be able to observe an effect of context at a later region of longer sentences. Another explanation that I provided is that the lexical bias of an animate subject limits the effect of contextual constraints during processing. This issue can be clarified by a future

[1]I am grateful to Professor Maria Piñango for this suggestion.

study that examines the configuration with an inanimate subject to compare with the AspVanimate configuration. Still another possible reason is that the context manipulation of the current study is too weak to constrain the readings effective enough to be observed in real-time measurements. This can be tested by a future experiment that further manipulates the context sentences to facilitate the biased reading more effectively.

Second, the experiments that examine the context effect on ambiguity resolution (Study 3 & 4) focus on the agentive along the eventive dimension vs. the constitutive reading along other dimensions. The constitutive reading of AspV sentences is still ambiguous among multiple dimensions (e.g., spatial, temporal, informational). While Study 4 shows that the ambiguity between the agentive and constitutive readings is resolved by biasing contexts, it remains unclear whether comprehenders disambiguate among different dimensions in the constitutive reading in a particular context. This can be examined by a different experimental design which pairs AspV sentences that contain inanimate subjects with different contexts that bias towards, for example, spatial vs. informational dimensions.

Third, this dissertation targets AspV sentences in its transitive use only. However, aspectual verbs can be used intransitively as well, as in "*The paper begins with an interesting question.*" In this intransitive use, the structured individual required by aspectual verbs is denoted by the subject rather than the complement. In addition, while the transitive use of AspVanimate sen-

tences licenses both the agentive and constitutive readings, the intransitive use seems to allow the constitutive reading only. From my own experiences, native speakers often point out that the intransitive use sounds more natural to express the constitutive reading. This is unsurprising because speakers may tend to use an unambiguous form to achieve clarity in communication. This said, how the intransitive use of aspectual verbs is processed in real-time remains a question. I conjecture that in the constitutive reading of AspV sentences, the transitive use (which licenses both agentive and constitutive readings) might be less acceptable than the intransitive use (which licenses the constitutive reading) due to the potential ambiguity in the former. As observed in Study 4, semantic ambiguity in sentences lowers acceptability.

Fourth, the semantic representations and real-time processing of psychological verbs have been discussed in literature yet remain at debate. The sentence interpretation task of Study 4 shows that sentences with psychological verbs (e.g., *enjoy*, *prefer*) give rise to both an eventive reading and a content reading. How these readings are processed (while the exact reading may remain underspecified) requires further investigations.

Finally, I have suggested in Chapter 9 that the architecture of the processing system implicated by the Structured Individual Hypothesis may be generalizable to other cases of meaning underspecification. In Chapter 2, I listed several coercion phenomena such as aspectual coercion. It will be interesting to see if the current processing view can account for the processing of

other coercion phenomena or cases involving meaning underspecification.

The issues mentioned above await to be addressed by future work, and I believe that the answers will shed light on the real-time comprehension of underspecified meaning.

Reference

Baayen, R Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Baayen, R Harald, Douglas J Davidson, and Douglas M Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language* 59:390–412.

Badre, David, Russell A Poldrack, E Juliana Paré-Blagoev, Rachel Z Insler, and Anthony D Wagner. 2005. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 47:907–918.

Baggio, Giosuè, Travis Choma, Michiel van Lambalgen, and Peter Hagoort. 2010. Coercion and compositionality. *Journal of Cognitive Neuroscience* 22:2131–2140.

Binder, Jeffrey R, Rutvik H Desai, William W Graves, and Lisa L Conant.

2009. Where is the semantic system? a critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex* 19:2767–2796.

Binder, Katherine S, and Keith Rayner. 1998. Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin & Review* 5:271–276.

Braze, David. 1999. Garden path sentences and morphological frequency biases. *University of Connecticut Working Papers in Linguistics, v10. Storrs, CT.* Cranberry Linguistics:pp. 13–34.

Brennan, Jonathan, and Liina Pylkkänen. 2010. Processing psych verbs: Behavioural and meg measures of two different types of semantic complexity. *Language and Cognitive Processes* 25:777–807.

Briscoe, Ted, Ann Copestake, and Bran Boguraev. 1990. Enjoy the paper: Lexical semantics via lexicology. In *Proceedings of the 13th conference on Computational linguistics-Volume 2*, 42–47. Association for Computational Linguistics.

Carpenter, Patricia A, and Marcel Adam Just. 1983. What your eyes do while your mind is reading. *Eye movements in reading: Perceptual and language processes* 275–307.

Cavanna, Andrea E, and Michael R Trimble. 2006. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129:564–583.

REFERENCE

Clifton, Charles, Matthew J Traxler, Mohamed Taha Mohamed, Rihana S Williams, Robin K Morris, and Keith Rayner. 2003. The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language* 49:317–334.

Crain, Stephen, and Mark Steedman. 1985. On not being led up the garden path: The use of context by the psychological parser. *Natural language parsing* 320–358.

Croft, William. 2012. *Verbs: Aspect and causal structure*. OUP Oxford.

Crozier, Sophie, Angela Sirigu, Stéphane Lehéricy, Pierre-François van de Moortele, Bernard Pillon, Jordan Grafman, Yves Agid, Bruno Dubois, and Denis LeBihan. 1999. Distinct prefrontal activations in processing sequence at the sentence and script level: An fmri study. *Neuropsychologia* 37:1469–1476.

Culicover, Peter W, and Ray Jackendoff. 2005. *Simpler syntax*. Oxford University Press Oxford.

Damasio, Hanna, Thomas J Grabowski, Daniel Tranel, Richard D Hichwa, and Antonio R Damasio. 1996. A neural basis for lexical retrieval. *Nature* 380:499.

De Almeida, Roberto G. 2004. The effect of context on the procesing of type-shifting verbs. *Brain and Language* 90:249–261.

REFERENCE

De Almeida, Roberto G, and Veena D Dwivedi. 2008. Coercion without lexical decomposition: Type-shifting effects revisited. *The Canadian Journal of Linguistics/La revue canadienne de linguistique* 53:301–326.

Delogu, Francesca, and Matthew Crocker. 2012. When is coercion simply surprisal? In *AMLaP 2012 Conference*, 211.

Delogu, Francesca, Matthew W Crocker, and Heiner Drenhaus. 2017. Teasing apart coercion and surprisal: Evidence from eye-movements and erps. *Cognition* 161:46–59.

Delogu, Francesca, Heiner Drenhaus, and Matthew Crocker. 2013. Teasing apart coercion and surprisal: evidence from ERPs and eye-movements. In *Talk given at the 26th Annual CUNY Conference on Human Sentence Processing*. Citeseer.

DiNardo, Laura. 2015. Competing analyses of complement coercion: new evidence from behavioral and electropsychophysiological methods. Senior thesis.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 547–619.

Duffy, Susan A, Robin K Morris, and Keith Rayner. 1988. Lexical ambiguity and fixation times in reading. *Journal of memory and language* 27:429–446.

REFERENCE

Evans, Alan C, D Louis Collins, SR Mills, ED Brown, RL Kelly, and Terry M Peters. 1993. 3d statistical neuroanatomical models from 305 mri volumes. In *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.*, 1813–1817. IEEE.

Ferreira, Fernanda, and Charles Clifton. 1986. The independence of syntactic processing. *Journal of memory and language* 25:348–368.

Fletcher, PC, CD Frith, SC Baker, T Shallice, RSJ Frackowiak, and RJ Dolan. 1995. The mind's eye—precuneus activation in memory-related imagery. *Neuroimage* 2:195–200.

Fodor, Jerry A., Thomas G. Bever, and Merrill Garrett. 1974. *The psychology of language: an introduction to psycholinguistics and generative grammar*. New York; Montreal: McGraw-Hill.

Fodor, Jerry A, and Ernie Lepore. 1998. The emptiness of the lexicon: reflections on james pustejovsky's the generative lexicon. *Linguistic Inquiry* 29:269–288.

Frazier, Lyn. 1979. *On comprehending sentences: Syntactic parsing strategies*. Bloomington, IN: Indiana University Linguistics Club.

Frazier, Lyn. 1987. Sentence processing: A tutorial review. In *Attention and performance 12: The psychology of reading*, ed. Max Coltheart, pp. 559-586. Lawrence Erlbaum Associates, Inc.

REFERENCE

Frazier, Lyn, and Janet Dean Fodor. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6:291–325.

Frazier, Lyn, and Keith Rayner. 1990. Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language* 29:181–200.

Frege, Gottlob. 1892. On concept and object. *The Frege Reader* 181–193.

Frisson, Steven, and Brian McElree. 2008. Complement coercion is not modulated by competition: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34:1.

Hickok, Gregory, and David Poeppel. 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92:67–99.

Hickok, Gregory, and David Poeppel. 2007. The cortical organization of speech processing. *Nature Reviews Neuroscience* 8:393–402.

Holmes, Colin J, Rick Hoge, Louis Collins, Roger Woods, Arthur W Toga, and Alan C Evans. 1998. Enhancement of mr images using registration for signal averaging. *Journal of computer assisted tomography* 22:324–333.

Humphries, Colin, Jeffrey R Binder, David A Medler, and Einat Liebenthal. 2007. Time course of semantic processes during sentence comprehension: an fmri study. *Neuroimage* 36:924–932.

REFERENCE

Husband, E. Mattew, Lisa A. Kelly, and David C. Zhu. 2011. Using complement coercion to understand the neural basis of semantic composition: evidence from an fMRI study. *Journal of Cognitive Neuroscience* 23:3254–3266.

Jackendoff, Ray. 1991. Parts and boundaries. *Cognition* 41:9–45.

Jackendoff, Ray. 1992. *Semantic structures*, volume 18. MIT press.

Jackendoff, Ray. 1997. *The architecture of the language faculty*. 28. MIT Press.

Jackendoff, Ray. 2002. *Foundations of language: brain, meaning, grammar, evolution*. Oxford University Press.

Jackendoff, Ray. 2007. A parallel architecture perspective on language processing. *Brain research* 1146:2–22.

Jackendoff, Ray. 2010. *Meaning and the lexicon: The parallel architecture 1975-2010*. OUP Oxford.

Katsika, Argyro, David Braze, Ashwini Deo, and Maria Mercedes Piñango. 2012. Complement coercion: distinguishing between type-shifting and pragmatic inferencing. *The Mental Lexicon* 7:58–76.

Krain, Amy L, Amanda M Wilson, Robert Arbuckle, F Xavier Castellanos, and Michael P Milham. 2006. Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making. *Neuroimage* 32:477–484.

REFERENCE

Krifka, Manfred. 1998. The origins of telicity. *Events and grammar* 197235.

Kuperberg, Gina R, Arim Choi, Neil Cohn, Martin Paczynski, and Ray Jackendoff. 2010a. Electrophysiological correlates of complement coercion. *Journal of Cognitive Neuroscience* 22:2685–2701.

Kuperberg, Gina R, Arim Choi, Neil Cohn, Martin Paczynski, and Ray Jackendoff. 2010b. Electrophysiological correlates of complement coercion. *Journal of Cognitive Neuroscience* 22:2685–2701.

Lapata, Mirella, Frank Keller, and Christoph Scheepers. 2003. Intra-sentential context effects on the interpretation of logical metonymy. *Cognitive Science* 27:649–668.

Lau, Ellen F, Colin Phillips, and David Poeppel. 2008. A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience* 9:920–933.

Levin, Beth, and Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge University Press.

Lewis, Shevaun, and Colin Phillips. 2015. Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research* 44:27–46.

MacDonald, Maryellen C, Neal J Pearlmutter, and Mark S Seidenberg. 1994a. The lexical nature of syntactic ambiguity resolution. *Psychological review* 101:676.

REFERENCE

MacDonald, Maryellen C, Neal J Pearlmutter, and Mark S Seidenberg. 1994b. Syntactic ambiguity resolution as lexical ambiguity resolution. In *Perspectives on sentence processing*, ed. Jr. Clifton, Charles, Lyn Frazier, and Keith Rayner, pp. 123–153. Lawrence Erlbaum Associates, Inc.

Magnani, Barbara, Francesca Frassinetti, Thomas Ditye, Massimiliano Oliveri, Marcello Costantini, and Vincent Walsh. 2014. Left insular cortex and left sfg underlie prismatic adaptation effects on time perception: Evidence from fmri. *NeuroImage* 92:340–348.

McElree, Brian, Liina Pylkkänen, Martin J. Pickering, and Matthew J. Traxler. 2006. A time course analysis of enriched composition. *Psychonomic Bulletin & Review* 13:53–59.

McElree, Brian, Matthew J. Traxler, Martin J. Pickering, Rachel E. Seely, and Ray Jackendoff. 2001. Reading time evidence for enriched composition. *Cognition* 78:B17–B25.

McRae, Ken, Michael J Spivey-Knowlton, and Michael K Tanenhaus. 1998. Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language* 38:283–312.

Moens, Mark, and Mark Steedman. 1987. Temporal ontology in natural language. In *Proceedings of the 25th annual meeting on Association for Computational Linguistics*, 1–7. Association for Computational Linguistics.

REFERENCE

Montague, Richard. 1973. *The proper treatment of quantification in ordinary English*. Springer.

Nolfe, E XMedCon, T Voet, F Jacobs, R Dierckx, and I XMedCon Lemahieu. 2003. An open-source medical image conversion toolkit. *Eur J Nucl Med* 30:S246.

Nunberg, Geoffrey. 1979. The non-uniqueness of semantic solutions: Polysemy. *Linguistics and philosophy* 3:143–184.

Pacht, Jeremy M, and Keith Rayner. 1993. The processing of homophonic homographs during reading: Evidence from eye movement studies. *Journal of Psycholinguistic Research* 22:251–271.

Papademetris, Xenophon, Marcel P Jackowski, Nallakkandi Rajeevan, Marcello DiStasio, Hirohito Okuda, R Todd Constable, and Lawrence H Staib. 2006. Bioimage suite: An integrated medical image analysis suite: An update. *The insight journal* 2006:209.

Partee, Barbara. 1987. Noun phrase interpretation and type-shifting principles. *Studies in discourse representation theory and the theory of generalized quantifiers.* 8:115–143.

Partee, Barbara, and Mats Rooth. 1983. Generalized conjunction and type ambiguity. *Formal Semantics: The Essential Readings* 334–356.

REFERENCE

Pearlmutter, Neal J, Kim G Daugherty, Maryellen C MacDonald, and Mark S Seidenberg. 1994. Modeling the use of frequency and contextual biases in sentence processing. *Urbana* 51:61801.

Phillips, Colin. 2013. Parser-grammar relations: We don't understand everything twice. *Language Down the Garden Path: The Cognitive and Biological Basis for Linguistic Structures* 294–315.

Pickering, Martin J., Brian McElree, and Matthew J. Traxler. 2005. The difficulty of coercion: A response to de Almeida. *Brain and Language* 93:1–9.

Piñango, Maria Mercedes. to appear. Concept composition during language processing: A model and two case studies .

Piñango, Maria Mercedes, and Ashwini Deo. 2012. Aspectual verbs and the "coercion" effect. In *Proceedings of Semantics and Linguistic Theory*, volume 22.

Piñango, Maria Mercedes, and Ashwini Deo. 2015. Reanalyzing the complement coercion effect through a generalized lexical semantics for aspectual verbs. *Journal of Semantics* ffv003.

Piñango, Maria Mercedes, and Edgar B. Zurif. 2001. Semantic operations in aphasic comprehension: implications for the cortical organization of language. *Brain and Language* 79:297–308.

REFERENCE

Pustejovsky, James. 1989. Current issues in computational lexical semantics. In *Proceedings of 4th Eur. ACL., Manchester,*, xvii–xxv.

Pustejovsky, James. 1991. The generative lexicon. *Computational linguistics* 17:409–441.

Pustejovsky, James. 1993. Type coercion and lexical selection. In *Semantics and the lexicon*, 73–94. Springer.

Pustejovsky, James. 1995. *The generative lexicon*. MIT Press, Cambridge, MA.

Pustejovsky, James, and Pierrette Bouillon. 1995. Aspectual coercion and logical polysemy. *Journal of semantics* 12:133–162.

Pustejovsky, James, and Elisabetta Jezek. 2008. Semantic coercion in language: Beyond distributional analysis. *Italian Journal of Linguistics* 20:175–208.

Pylkkänen, Liina. 2008. Mismatching meanings in brain and behavior. *Language and Linguistics Compass* 2:712–738.

Pylkkänen, Liina, Andrea E. Martin, Brian McElree, and Andrew Smart. 2009. The anterior midline field: coercion of decision making? *Brain and Language* 108:184–190.

Pylkkänen, Liinas, and Brian McElree. 2006. The syntax-semantics interface:

on-line composition of sentence meaning. In *Handbook of psycholinguistics*, ed. Morton A. Traxler, Matthew J; Gernsbacher. New York: Elsevier.

Pylkkänen, Liinas, and Brian McElree. 2007. An MEG study of silent meaning. *Journal of Cognitive Neuroscience* .

R Core Team. 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org/`.

R Core Team. 2015. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL `https://www.R-project.org/`.

Rayner, Keith, and Lyn Frazier. 1989. Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:779.

Rodd, Jennifer M, Olivia A Longe, Billi Randall, and Lorraine K Tyler. 2010. The functional organisation of the fronto-temporal language system: evidence from syntactic and semantic ambiguity. *Neuropsychologia* 48:1324–1335.

Rozwadowska, Bożena. 1989. Are thematic relations discrete. *Linguistic categorization* 115–130.

REFERENCE

Seidenberg, Mark S, Michael K Tanenhaus, James M Leiman, and Marie Bienkowski. 1982. Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive psychology* 14:489–537.

Sereno, Sara C, Patrick J O'Donnell, and Keith Rayner. 2006. Eye movements and lexical ambiguity resolution: Investigating the subordinate-bias effect. *Journal of Experimental Psychology: Human Perception and Performance* 32:335.

Shapiro, Lewis P, Betty Gordon, Nancy Hack, and Jennifer Killackey. 1993. Verb-argument structure processing in complex sentences in broca s and wernicke s aphasia. *Brain and language* 45:423–447.

Shapiro, Lewis P, Edgar B Zurif, and Jane Grimshaw. 1989. Verb processing during sentence comprehension: Contextual impenetrability. *Journal of psycholinguistic research* 18:223–243.

Shetreet, Einat, Tal Linzen, and Naama Friedmann. 2016. Against all odds: exhaustive activation in lexical access of verb complementation options. *Language, Cognition and Neuroscience* 31:1206–1214.

Shetreet, Einat, Dafna Palti, Naama Friedmann, and Uri Hadar. 2007. Cortical representation of verb processing in sentence comprehension: Number

of complements, subcategorization, and thematic frames. *Cerebral Cortex* 17:1958–1969.

Spivey, Michael J, and Michael K Tanenhaus. 1998. Syntactic ambiguity resolution in discourse: modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24:1521.

Stadler, Waltraud, Ricarda I Schubotz, D Yves von Cramon, Anne Springer, Markus Graf, and Wolfgang Prinz. 2011. Predicting and memorizing observed action: differential premotor cortex involvement. *Human brain mapping* 32:677–687.

Studholme, Colin, Edward Novotny, IG Zubal, and JS Duncan. 2001. Estimating tissue deformation between functional images induced by intracranial electrode implantation using anatomical mri. *Neuroimage* 13:561–576.

Swinney, David A. 1979. Lexical access during sentence comprehension:(re) consideration of context effects. *Journal of verbal learning and verbal behavior* 18:645–659.

Talmy, Leonard. 1988. Force dynamics in language and cognition. *Cognitive science* 12:49–100.

Tanaka, Satoshi, Manabu Honda, and Norihiro Sadato. 2005. Modality-specific

cognitive function of medial and lateral human brodmann area 6. *Journal of Neuroscience* 25:496–501.

Tanenhaus, Michael K, Susan M Garnsey, and Julie Boland. 1990. Combinatory lexical information and language comprehension. In *Cognitive models of speech processing: Psycholinguistic and computational perspectives.*, ed. Gerry T. M Altmann, pp. 383–408. The MIT Press.

Tanenhaus, Michael K, James M Leiman, and Mark S Seidenberg. 1979. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of verbal learning and verbal behavior* 18:427–440.

Tanenhaus, Michael K, Michael J Spivey-Knowlton, and Joy E Hanna. 2000. Modeling thematic and discourse context effects with a multiple constraints approach: Implications for the architecture of the language comprehension system. *Architectures and mechanisms for language processing* 90–118.

Thompson, Cynthia K, KL Lange, Sandra L Schneider, and Lewis P Shapiro. 1997. Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology* 11:473–490.

Traxler, Matthew J., Rihana S. McElree, Brian and; Williams, and Martin J. Pickering. 2005. Context effects in coercion: evidence from eye movements. *Journal of memory and language* 53:1–25.

Traxler, Matthew J, Martin J Pickering, and Brian McElree. 2002. Coercion in

sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language* 47:530–547.

Trueswell, J, and M Tanenhaus. 1994. Toward a lexical framework of constraint-based syntactic ambiguity resolution. *Perspectives on sentence processing* 155–179.

Trueswell, John C. 1996. The role of lexical frequency in syntactic ambiguity resolution. *Journal of memory and language* 35:566–585.

Utt, Jason, Alessandro Lenci, Sebastian Padó, Alessandra Zarcone, et al. 2013. The curious case of metonymic verbs: A distributional characterization. In *Proceedings of the IWCS 2013 Workshop Towards a Formal Distributional Semantics*.

Van Gompel, Roger PG, Martin J Pickering, and Matthew J Traxler. 2001. Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language* 45:225–258.

Verkuyl, Henk, and Joost Zwarts. 1992. Time and space in conceptual and logical semantics: the notion of path. *Linguistics* 30:483–512.

Wallentin, Mikkel, Ethan Weed, Leif Østergaard, Kim Mouridsen, and Andreas Roepstorff. 2008. Accessing the mental space—spatial working memory processes for language and vision overlap in precuneus. *Human Brain Mapping* 29:524–532.