LOCATIVE VERBS IN L2 LEARNING: A MODULAR PROCESSING PERSPECTIVE

By

Roman Chepyshko

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Second Language Studies – Doctor of Philosophy

2018
ABSTRACT

LOCATIVE VERBS IN L2 LEARNING: A MODULAR PROCESsing PERSPECTIVE

By

Roman Chepyshko

The current project investigates developmental aspects of acquiring locative verb constructions in English as a second language. Locative verbs, such as to pour, to spill, to spray, and to sprinkle, constitute a prototypical case of an overgeneralization problem in language learning: Whereas some of the verbs can appear in two alternative constructions (e.g., John sprayed water onto the plant vs. John sprayed the plant with water), others are restricted to a single frame (e.g., John poured juice into the glass) and are perceived as unnatural when used in the alternative structure (e.g., *John poured the glass with juice). How do language learners acquire the alternation patterns, and, at the same time, avoid overgeneralizations? In contrast to the depth and diversity of child language investigations (e.g., Ambridge, Pine, & Rowland, 2012; Gropen, Pinker, Hollander, & Goldberg, 1991a, 1991b), only a handful of research projects examined acquisition of the locative semi-regularities in L2 producing contradictory and largely irreconcilable proposals. The dissertation critically reviews methodological, theoretical, and empirical underpinnings of the conflicting claims, recasts the problem of learning locative semi-regularities within the MOGUL (Modular Online Growth and Use of Language) framework (Sharwood Smith & Truscott, 2014; Truscott & Sharwood Smith, 2004), and proposes an alternative processing-based account of how real-time encounters of the target constructions relate to the developmental transitions in language learners’ grammatical representations.
The study tested predictions of the alternative proposals in two empirical investigations. One hundred and two native speakers of Mandarin Chinese learning English as a second language (of low, intermediate, and high L2 proficiency levels) and a comparison group of English native speakers (n=50) took part in an acceptability judgment task (Experiment 1) and an eye-tracking study using visual world paradigm (Experiment 2). Results revealed complex but consistent patterns in participants’ performance. All experimental groups demonstrated both target-like performance and diverged from the assumed target language standards. These findings suggest that any property theory relying on specifications of what L2 learners know (or fail to know) at different stages in L2 development (Gregg, 2008) would face considerable challenges in explaining the observed behavioral patterns. At the same time, the results can be naturally understood within a dynamic framework allowing an integration of the property-theoretical constructs with the mechanisms explaining real-time changes in the linguistic system.
ACKNOWLEDGEMENTS

There are many people to whom I am grateful for their support and encouragement in the development of the current dissertation. I am thankful to Bill VanPatten for his guidance, involvement, and encouragement throughout the years of my graduate studies. My deepest gratitude goes to Paula Winke for her insightful feedback and inspirational comments. I am thankful to Debra Hardison who provided her expertise to this dissertation and offered thought-provoking comments. My special thanks go to Aline Godfroid for the years of true involvement, encouragement, and support. I am also extremely grateful to Susan Gass for her leadership in the Second Language Studies program, incredible capacity to hear and support, as well as her feedback and comments throughout my work on the dissertation project.

I would like to thank my professors at Michigan State University, Diogo Almeida, Eric Altmann, Alan Beretta, Karthik Durvasula, Debra Friedman, Shawn Loewen, and Patti Spinner. I greatly appreciate your teaching and value your professionalism and dedication. Also, I would like to thank John Truscott, Mike Sharwood Smith, and Johanna Katchen for supporting me in pursuing the doctoral studies. Many thanks to my colleagues and friends, Jens, Scott, Se Hoon, Ayman, Dominik, Le Anne, Jimin, Ina, and Ji-Hyun.

My special thanks go to the Greys, Jeremy, Heather, Dylan, and Nigel. Thank you for all the incredible things. I also would like to thank my family. My parents-in-law, Yuriy and Olga Stetsko, and my brother-in-law, Pavlo Stetsko, thank you for your encouragement and prayers throughout the years. I am grateful to my wife, Hanna Chepyshko, for her incredible patience, encouragement, and believing in me. Thanks mom for your endless love and care. Finally, I would like to thank my daughter Solomia and my son Marko. I love you.
# TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................................................... viii

LIST OF FIGURES ........................................................................................................................................... x

INTRODUCTION .............................................................................................................................................. 1

CHAPTER 1: LITERATURE REVIEW .................................................................................................................. 5
   Semantics of the locative verbs ..................................................................................................................... 5
   Syntax-semantic correspondences of locative verbs ..................................................................................... 7
   Acquisition of argument structure of locative verbs ................................................................................... 8
       Pinker’s semantic category account ......................................................................................................... 9
       Usage-based perspective on the retreat from overgeneralization ............................................................ 12
       Integrating semantics and statistics: A hybrid account ........................................................................... 14
   Locative constructions in L2 acquisition ..................................................................................................... 16
       Juffs (1996) ............................................................................................................................................. 17
       Bley-Vroman and Joo (2001) ................................................................................................................... 19
       Joo (2003) ............................................................................................................................................. 21
       Ambridge and Brandt (2013) .................................................................................................................. 22
   Critical review ........................................................................................................................................... 24
   Methodological sources of the conflicting accounts .................................................................................... 29
   Conclusion ................................................................................................................................................ 32
   A processing perspective on learning locative constructions .................................................................... 34
   Modular Online Growth and Use of Language (MOGUL) .......................................................................... 36
   Acquisition of locative verbs’ argument structure ...................................................................................... 41
   Summary .................................................................................................................................................. 48
   The current empirical study ....................................................................................................................... 49

CHAPTER 2: EXPERIMENT 1 .............................................................................................................................. 51
   Method ..................................................................................................................................................... 51
   Participants ............................................................................................................................................... 54
       Native speakers of English ..................................................................................................................... 54
       L2 speakers of English ........................................................................................................................... 54
   Materials ................................................................................................................................................ 56
       Sentence Stimuli ..................................................................................................................................... 56
       Acceptability Judgments Questionnaire ................................................................................................. 58
       Background questionnaire ..................................................................................................................... 59
       Proficiency measurement instrument .................................................................................................. 59
       Vocabulary Knowledge Scale .............................................................................................................. 60
   Procedures .............................................................................................................................................. 60
   Data processing ........................................................................................................................................ 61
CHAPTER 3: EXPERIMENT 2

Method ..........................................................................................................................85
  Participants ..................................................................................................................87
  Materials .......................................................................................................................87
    Sentence stimuli .........................................................................................................87
    Auditory and visual stimuli .......................................................................................89
    Experimental presentations ......................................................................................90
  Procedure .....................................................................................................................91
  Data processing ...........................................................................................................92
  Statistical Analysis ......................................................................................................93
 Results ............................................................................................................................96
  Overview of the eye moment patterns ..................................................................96
  Time course analysis .................................................................................................99
    English native speakers .........................................................................................99
    High L2 group .........................................................................................................102
    Intermediate L2 group ...........................................................................................103
    Low L2 group ..........................................................................................................106
  Summary of the growth curve analysis ................................................................108
  Examining post-verbal predictions ......................................................................108
    English native speakers .........................................................................................109
    High L2 group ..........................................................................................................110
    Intermediate L2 group ............................................................................................112
    Low L2 group ..........................................................................................................113
  Summary of the experimental results ..................................................................115

CHAPTER 4: DISCUSSION .........................................................................................119
LIST OF TABLES

Table 1. Summary of empirical studies investigating locative verb constructions in L2 ............24
Table 2. The demographic information for Low, Intermediate, and High proficiency L2 groups 55
Table 3. The target locative verbs used in the study .......................................................56
Table 4. Percentage of accepting non-alternating constructions by the experimental groups .........83
Table 5. The non-alternating locative verbs and two sets of content-container arguments ........88
Table 6. The mean duration and standard deviation (in parenthesis) in milliseconds for verb, article, and noun segment of the target auditory stimuli by verb type ..................................90
Table 7. Summary of growth curve regression results for content and container non-alternating sentences for L1 English group ........................................................................100
Table 8. Summary of growth curve regression results for content and container non-alternating sentences for High L2 group ........................................................................103
Table 9. Summary of growth curve regression results for content and container non-alternating sentences for Intermediate L2 group .................................................................105
Table 10. Summary of growth curve regression results for content and container non-alternating sentences for Low L2 group ..................................................................................107
Table 11. Odds ratio of looking at Container images in processing container non-alternating locative constructions by experimental groups .........................................................117
Table 12. Means and standard deviations in parenthesis of locative verbs’ characteristics ....157
Table 13. Means and standard deviations for timings of participants’ responses in Experiment 1 .........................................................................................................................................165
Table 14. Percentage of discarded test items in Experiment 1 arranged by experimental groups ........................................................................................................................................166
Table 15. Percentage of “uncertain” responses in Experiment 1 organized by proficiency groups .......................................................................................................................................167
Table 16. Summary of logistic regression model for the content non-alternating locative stimuli .................................................................................................................................169
Table 17. Summary of logistic regression model for container non-alternating locative stimuli .............................................................................................................................................169
Table 18. *Summary of logistic regression model for the content alternating locative stimuli*....170
Table 19. *Summary of logistic regression model for the container alternating locative stimuli* 170
Table 20. *Mixed-effect model summary for non-alternating locatives’ preference scores*........171
Table 21. *Summary of the grow curve regression results for L1 English group*......................176
Table 22. *Summary of the grow curve regression results for High L2 group*..........................176
Table 23. *Summary of the grow curve regression results for Intermediate L2 group*.............177
Table 24. *Summary of the grow curve regression results for Low L2 group*..........................177
LIST OF FIGURES

Figure 1. Functional architecture of language in MOGUL.........................................................37

Figure 2. “Uniform differences” (preference for content over container-oriented uses of the locative verbs by categories) for L1 (x axis) and Low L2 (y axis) speakers of English..............67

Figure 3. “Uniform differences” (preference for content over container-oriented uses of the locative verbs by categories) for L1 (x axis) and Intermediate L2 (y axis) speakers of English. 68

Figure 4. “Uniform differences” (preference for content over container-oriented uses of the locative verbs by categories) for L1 (x axis) and High L2 (y axis) speakers of English..............69

Figure 5. Mean acceptance rate of the container-alternating locative verbs by the experimental groups. Error bars represent standard errors.................................................................73

Figure 6. Mean acceptance rate of content-alternating locative verbs by the experimental groups. Error bars represent standard errors.................................................................73

Figure 7. Mean acceptance rate of the container non-alternating locative verbs by the experimental groups. Error bars represent standard errors.................................................................74

Figure 8. Mean acceptance rate of the content non-alternating (right) locative verbs by the experimental groups. Error bars represent standard errors.................................................................74

Figure 9. Mean preference scores for content and container non-alternating locative verbs for low, intermediate and high proficiency L2 speakers. Error bars represent standard errors.............79

Figure 10. Effect of lexical frequency on High L2 participants’ ratings of content non-alternating locative verbs in content-oriented (blue dots) and container-oriented (red dots) sentence frames.................................................................82

Figure 11. Time course graph for L1 English group showing likelihood of fixations to the Content (black circles) and Container (white triangles) visual images viewed during auditory presentation of Content-only (left panel) and Container-only locative sentences.........................97

Figure 12. Time course graph for High L2 English group showing likelihood of fixations to the Content (black circles) and Container (white triangles) visual images viewed during auditory presentation of Content-only (left panel) and Container-only locative sentences.........................97

Figure 13. Time course graph for Intermediate L2 group showing likelihood of fixations to the Content (black circles) and Container (white triangles) visual images viewed during auditory presentation of Content-only (left panel) and Container-only locative sentences.........................98

x
Figure 14. Time course graph for Low L2 group showing likelihood of fixations to the Content (black circles) and Container (white triangles) visual images viewed during auditory presentation of Content-only (left panel) and Container-only locative sentences...........................................98

Figure 15. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the verb segment of content-only (left panel) and container- only locative sentences for L1 English group.................................................................99

Figure 16. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the verb segment of content-only (left panel) and container- only locative sentences for High L2 group.................................................................102

Figure 17. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the verb segment of content-only (left panel) and container- only locative sentences for Intermediate L2 group.........................................................104

Figure 18. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the verb segment of content-only (left panel) and container- only locative sentences for Low L2 group.................................................................107

Figure 19. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the direct object segment of content-only (left panel) and container- only locative sentences for L1 English group.........................................................109

Figure 20. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the direct object segment of content-only (left panel) and container- only locative sentences for High L2 group.................................................................111

Figure 21. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the direct object segment of content-only (left panel) and container- only locative sentences for Intermediate L2 group.........................................................112

Figure 22. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the direct object segment of content-only (left panel) and container- only locative sentences for Low L2 group.................................................................114
INTRODUCTION

Language learning is woven from multiple instances in which learners encounter small bits of the linguistic material. Through hearing stretches of patterned sounds, young children learning their first language (L1) as well as adult second language (L2) learners acquire complex generalizations which should be unlearnable from fractured input (e.g., Hauser, Chomsky, & Fitch, 2002; Pinker & Jackendoff, 2005). This generalization capacity is a key feature of human cognition. It allows for predictions of unknown properties of the surrounding world on the basis of inherently limited experiential evidence (e.g., Chang, Kidd, & Rowland, 2013; Jaeger & Ferreira, 2013; Johnson, Turk-Browne, & Goldberg, 2013; Pickering & Garrod, 2007, 2013). On the downside, such predictions can often turn out to be erroneous and nonconforming with the actual state of affairs. In language acquisition, stepping beyond the observed evidence may lead to a range of overgeneralizations resulting in ungrammaticalities of a target language.

Overgeneralizations are a common part of language development. They reflect language acquirers’ tendencies to make regular irregularities of the language they learn (e.g., Hudson Kam & Chang, 2010; Hudson Kam & Newport, 2005, 2009; Perfors, 2012). On many occasions, such overgeneralizations are fully consistent with the available evidence, raising the question of how language learners manage to constrain their overly general grammars given an apparent arbitrariness of the exceptions and no environmental evidence restricting their productive patterns.

The current project investigates the developmental aspects of acquiring locative verb constructions in the L2 (Brinkmann, 1997; Iwata, 2005, 2008, Pinker, 2008, 2009). Locative verbs, such as to pour, to spill, to spray, and to sprinkle, constitute a prototypical case of the partial generalization problem in language learning: Whereas some of the verbs can appear in
two alternative constructions (e.g., *John sprayed water onto the plant vs. John sprayed the plant with water*), others are restricted to a single frame (e.g., *John poured juice into the glass*) and are perceived as ungrammatical when used in the alternative structure (e.g., *John poured the glass with juice*). How do L2 learners deal with the task of learning locative constructions? Are they able to acquire the alternation patterns, and, at the same time, to restrict their generalizations to an apparently arbitrary subset of the locative verbs? To what extent do L2 learners’ native language, their L2 experiences, and language learning predispositions determine success or failure in the development of the target constructions? Do L2 learners follow the developmental path of L1 acquirers or take an alternative route in the acquisition of the semi-regularities? The payoff of addressing these questions is far beyond explaining acquisition of a specific grammatical construction in L2. An understanding of the semi-regularities is an essential step towards an understanding of the relationship between L2 learners’ linguistic experiences and their L2 knowledge.

In contrast to the depth and diversity of L1 investigations (e.g., Ambridge, Pine, & Rowland, 2012; Bidgood, Ambridge, Pine, & Rowland, 2014; Brinkmann, 1997; Gropen, Pinker, Hollander, & Goldberg, 1991a, 1991b; Kim, Landau, & Phillips, 1990; Twomey, Chang, & Ambridge, 2014), only a handful of research projects examined acquisition of the locative semi-regularities in L2 producing contradictory and largely irreconcilable proposals. Juffs (1996) explained growth of the locative constructions in L2 in terms of cross-linguistic transfer governed by lexical parameter settings of L2 learners’ L1. Bley-Vroman and Joo (2001) and Joo (2003) held that L2 learners in principle fail to recover from the overgeneralizations and cannot rely on their native language in the acquisition of the locatives. In turn, Ambridge and Brandt (2013) suggested that L2 learners do attain target-like knowledge of the locatives but do so in a
categorically different manner from L1 learners. Despite the apparent contradictions, none of the published investigations have attempted to reexamine the mutually exclusive claims, in effect, leaving the field in a rather stagnant position.

The proliferation of conflicting claims reflects a lack of common theoretical ground among the alternative proposals. The previous researchers approached the research topic with different sets of assumptions and, to a certain degree, tended to promote their initial positions even when empirical evidence did not support them. Each publication hypothesized distinctive properties of L2 learners’ grammars (i.e., states of their grammatical knowledge) and used different behavioral tasks and statistical procedures to validate presence/absence of such properties in participants’ experimental performance. At the same time, the authors of these studies left underspecified the processes by which L2 learners’ knowledge can be established in the first place: They said next to nothing about how (or why) instances of a target language heard or read by L2 learners enable or fail to enable changes in the learners’ mental grammars. In essence, these proposals abstracted away from the generalization part of the partial generalization problem.

The main goal of this dissertation is to revive the study of the locative constructions in L2 acquisition by resolving theoretical and empirical inconsistencies observed in the L2 literature. Chapter 1 deals with the theoretical aspects of the partial generalization problem. Here, I first present the key constructs and solutions to the problem of learning locative semi-regularities proposed within the L1 acquisition field, therefore building the necessary background for an understanding of the locative phenomena. Then, I review and critically reexamine L2 studies identifying theoretical and methodological sources of the conflicting claims. In the final part of Chapter 1, I recast the problem of learning locative semi-regularities within the MOGUL
(Modular Online Growth and Use of Language) framework (Sharwood Smith, 2017; Sharwood Smith & Truscott, 2014; Truscott, 2014; Truscott & Sharwood Smith, 2004) and propose an alternative processing-based account specifying how real-time encounters of the target constructions can be related to the developmental transitions in language learners’ grammatical representations. The theoretical reexamination of the locative phenomena is followed by two experimental studies in Chapters 2 and 3. The empirical investigations explore offline (Chapter 2) and online (Chapter 3) perception of the target constructions and are aimed at identifying patterns in L2 learners’ performance. As such, the experiments serve the goal of obtaining detailed and unbiased empirical evidence needed for the reassessment of the alternative proposals of the locative phenomena in L2 acquisition. Finally, in Chapter 4, I summarize the experimental results, discuss the validity of the existing alternatives, revise the processing-based account, and identify directions for future research.
CHAPTER 1: LITERATURE REVIEW

This chapter introduces the theoretical and empirical background pertaining to the phenomena of learning locative constructions. It is structured in the following way. The first part sets the stage by defining the semantic underpinning of the locative verbs and specifying how the semantic structures are realized in syntax. Then, it reviews major theoretical accounts on how the semi-regular syntax-semantics correspondences are established in L1 development. The second part focuses on research in L2. It offers a detailed overview of the published investigations and then critically examines each study identifying theoretical and methodological shortcomings of the conflicting proposals. The final part recasts the problem of learning the partially productive patterns in processing terms and proposes a processing-based solution integrating semantic and statistical mechanisms responsible for constraining the overproductivity.

Semantics of the locative verbs

Locative verbs denote motion events in which substances or objects (Content) move across a space into or onto a location (Container) (for an introduction of locative phenomena, see Arad, 2007). In most inclusive terms, these verbs signify the relationship between Containers and Contents caused or enabled by an Agent performing an action. For example, locative verbs in (1) all specify the manner in which the Contents (water, paint, grease, and coffee) and the Containers (glass, wall, axle, and jeans) interact as a result of the Agents’ involvement.

(1) a. The waiter filled the glass with water.
   b. Donna sprayed the paint on the wall.
   c. The mechanic smeared the axle with grease.
   d. John spilled coffee on his jeans.
The Content-Container relationship is not something given in the environment independently from an observer. Rather, it is the product of people’s conceptualization of the surrounding world: The process by which they make sense of the environment by assigning specific roles and imposing organization on the perceptual continuum. Importantly, the same arrangement of the entities and substances in space and time can be construed in different ways. For instance, both verbs to pour and to fill can be used to describe an action in which the Content moves in a downward continuous stream concurrently occupying the capacity of Container. In such case, for example, we can speak of either pouring tea into the cup or filling the cup with tea, accordingly adopting alternative perspectives on the observed event.

The semantic core of the verb to pour specifies caused motion of the Content (tea) into the Container (the cup) and can be schematically represented by (2a) (Levin, 1993; Pinker, 1989; Rappaport Hovav & Levin, 2005). Here, an Agent (X) acts upon a Content (Y), causing its transfer to a Container (Z). In contrast, the semantic structure underlying the verb to fill represents the change of state of the Container (the cup) by means of moving the Content (tea) into the receptacle. The representation in (2b) captures this generalization: an Agent (X) causes the state change of the Container (Z) by moving the Content (Y) to the Container (Z).

(2)  

a. X causes Y to go to Z  

b. X causes Z to change state by means of causing Y to go to Z

The caused motion of the Content (into the Container) and a change of state of the Container (arising from transferring the Content to the Container) are two alternative ways of construing locative events. Critically, they differ in which participant in the observed scene is thought to be primarily affected by an Agent’s action: either Content or Container. In reality, however, this distinction is quite subtle allowing alternative interpretations of any locative event.
By definition, any state change of the receptacles in a locative event involves caused motion of Content. Similarly, by adding stuff to a location (causing a Content’s motion to a Container), the speakers potentially affect the final destination. Therefore, when we communicate linguistically about locative events, people rely on structured conceptual representations, which, nevertheless, are highly amendable to the alternative interpretations.

**Syntax-semantic correspondences of locative verbs**

English locative verbs demonstrate a high degree of variability in how they express the locative arguments syntactically. A subset of the locative verbs, comprising, for example, *to pour, to drip, to dump,* and *to spill* exclusively encodes the Content argument in the direct object position. These types of locative verbs sound unnatural in the alternative constructions. These *content-only* locative verbs express caused-motion aspect of the locative events specifying how the affected entities are set in motion by an Agent or in what manner these objects move. For instance, the verbs *to dump* and *to spill* indicate whether the Content’s motion gets initiated unintentionally (e.g., *John spilled coffee…*) or in hurried and/or careless manner (e.g., *Mary dumped the papers…*). Similarly, the verbs *to pour* and *to drip* specify whether the substances move in small drops of liquid or in a steady downward stream (e.g., *The cook poured/dripped oil onto the pan*).

A second category of the locatives, such as *to fill, to cover, to stain,* and *to decorate* selectively realizes Containers as direct objects. The *container-only* locative verbs express the state change of Containers underspecifying the manner in which the Contents move or are caused to move to the receptacles. For instance, the acts of filling, covering, staining or decorating can be accomplished in any imaginable manner. What really matters is the change of the Containers’
states in either spatial or aesthetic terms: when the Containers become full, bear an additional layer of stuff, or change in their attractiveness.

The third group of the locatives, including to spray, to spread, and to pile are capable of accommodating either Content or Container arguments as their direct objects. In the content-oriented sentence constructions (Contents as direct objects) these verbs signify the manner of the substances’ motion. In the alternative container-oriented frames (Containers as direct objects), they indicate the ways in which the receptacles change their states. In technical terms, we can speak of these verbs as alternating between two different ways of argument expression. Therefore, we can speak of either spraying pesticides or spraying the tree, spreading tomato paste or spreading the toast, piling dishes or piling the table creatively shifting the meanings of the verbs.

**Acquisition of argument structure of locative verbs**

Learning to talk about locative events is essentially the task of acquiring partially productive generalizations: It necessarily involves acquisition of the regular alternation patterns alongside the seemingly arbitrary irregularities. Such a task is often described as a *learning paradox* stemming from a simultaneous assumption of the following premises: (a) productivity of linguistic generalizations, (b) arbitrariness of exceptions, and (c) lack of direct negative evidence informing learners about the exceptions to the productive patterns (Baker, 1979; Goldberg, 2006; Pinker, 2009). The productivity is the minds’ ability to generate novel constructions, which on many occasions involves stepping beyond learners’ experiential evidence. The apparent arbitrariness of the exceptions is best reflected by the following examples: The locative verbs to splash and to spray (splash/spray water onto the face vs. splash /spray the face with water) but not their semantic neighbors to pour and to drip (pour/drip water
onto the floor vs. *pour/*drip the floor with water) undergo the locative alternation. Even if there is something special about alternating splashing and spraying that distinguishes these words from non-alternating pouring and dripping, these distinctions are not obvious and appear to be purely arbitrary. Finally, the no negative evidence premise reflects an observation that the linguistic input is generally void of direct evidence signaling the ill-formed locative constructions. There is nothing in the ambient language diverting learners from applying the productive generalizations and, therefore, forming such sentences as *pour/*drip the floor with water. If all advanced premises hold true, the only possible logical conclusion is that the partially productive patterns are in principle unlearnable. The ability of predicting novel constructions (productivity) extended to the set of the unpredictable (arbitrary) cases should result in consistent ungrammaticalities with nothing in the ambient language indicating such cases as ill-formed (no negative evidence). Contrary to such a scenario, native speakers of English do manage to acquire the target constructions demonstrating the paradoxical learning.

Contemporary linguistic research has offered several notable solutions of the learning paradox. Besides explaining acquisition of specific grammatical constructions, these accounts have contributed to the long-standing question concerning the nature of linguistic knowledge: They channel the scholarly discourse towards in-depth examination of the linguistic environment, language learners’ experiences and learning predispositions, as well as their interplay in the course of language acquisition. In what follows, I present a review of major accounts advancing alternative perspectives on learning the partially productive generalizations.

Pinker’s semantic category account

The approach developed by Pinker (1989) relies on a learning process in which language learners progressively discover and refine lexical semantics of the locative verbs (also, see
Gropen et al., 1991a, 1991b; Pinker, 2008). As an initial step towards solving the paradox Pinker tentatively subdivided the locative verbs into two broad categories based on their primary semantics: those that primarily denote caused motion of Content and the ones signifying Container’s change of state. The membership in the broad classes determines the common argument structure of the verbs sharing the same semantics. The broad-range rules align specific construals of the locative events with their formal structures. The correspondences between a particular construal of the event and its expression in syntax are governed by innate linking rules specifying which conceptual participant appears in which syntactic position. Specifically, these rules link the causal agent of the event to the grammatical subject, the affected entity to the direct object, and the final destination of the content’s motion or the means of changing the container to the indirect object position. For example, if a verb denotes the manner in which an entity changes its location, the Content argument becomes syntactically realized as a direct object. Conversely, if a verb primarily specifies how a receptacle changes its state, the argument undergoing changes (i.e., Container) is delivered syntactically in the direct object position.

Given that each locative verb falls into one of the two broad categories, Pinker (1989) proposed that the alternative constructions are derived grammatical forms. The derivational rules operating on the lexical semantic structures govern transformation of one construal of a locative event into another. Their basic functioning can be described as a conceptual gestalt shift altering the content oriented semantic structures into container oriented and vice versa.

The alternation rules allow transforming one semantic frame into another successfully predicting the argument realization patterns of such verb as pack, load, splash, spray, and spread. The flexibility of such generalizations, however, comes at a cost. It falsely predicts that all locative verbs can appear in the alternative constructions generating ungrammatical structures
as *pour the glass with water or *fill water into the glass. To explain how children learn to avoid the overgeneralizations, Pinker proposes narrow-range rules constraining the alternation patterns. These rules are generalizations over sets of locative verbs sharing certain semantic properties. They organize semantically similar verbs into small subclasses or categories. The membership in these somewhat idiosyncratic semantic categories then determines the ability of a particular lexeme to alternate or not alternate. A subset of the narrow-range rules is presented in (3).

(3) Narrow-range semantic classes adopted from Pinker (1989, p. 273-274).

- **Smear** class (alternating) Simultaneous forceful contact and motion of a mass against a surface: brush, dab, plaster, rub, slather, smear, smudge, spread, and streak.
- **Pour** class (content non-alternating) A mass is enabled to move via the force of gravity: dribble, drip, drizzle, dump, ladle, pour, shake, slop, slosh, spill.
- **Fill** class (container non-alternating) A layer completely covers a surface: bandage, blanket, coat, cover, encrust, fill, flood, inlay, pave, smother, etc.

The narrow-range rules are composites of the semantic features capturing fine-grain distinctions between a family of causal relations (i.e., causing, letting and enabling) (Jackendoff, 1990a, 2003, 2006, Talmy, 1988, 2000), aspects of objects’ spatial geometry (Jackendoff, 1991, 1996), semantic properties specifying construal of matter as either unbounded mass, aggregate or a bounded/countable entity (Jackendoff, 1987b; Landau & Jackendoff, 1993). They are specific instantiations of the broad-range generalizations determining the alternation status of the locative verbs by the verbs’ compatibility with the broad semantic classes. For example, the locatives in the smear class denote actions in which an agent acting directly upon a substance simultaneously affects a surface of a receptacle. Therefore, these verbs are in principle compatible with either caused motion or state-change interpretations. In contrast, the verbs of the pour-type denote
activities in which an agent’s role is limited to enabling motion events with the intermediary force delivering content to its final destinations. The presence of the intermediary force explains the relative unpredictability of how substances in motion affect containers. Therefore, the members of this narrow-range class are compatible only with the caused motion interpretation and reserve their argument realization to a single syntactic frame.

In sum, Pinker advanced a well-articulated account of learning the partially productive generalizations. In essence, the author dismissed the arbitrariness of exceptions premise specifying the necessary (broad-range rules) and sufficient (narrow-range rules) conditions for the productive use of the locative alternation patterns. Thus, on this account the learning of the argument realization of the locative verbs proceeds over time as language learners acquire meanings of the locatives, refining the distinctions denoted by the verbs.

Usage-based perspective on the retreat from overgeneralization

The second account of how language learners manage to acquire the partially productive generalizations relies on the distributional properties of the locative verbs (e.g., Braine & Brooks, 1995; Brooks, Tomasello, Dodson, & Lewis, 1999; Goldberg, 2011, 2016; Robenalt & Goldberg, 2015; Theakston, 2004). In a nutshell, the proposed solution of the learning paradox maintains that language acquirers do have the access to the negative evidence available in the linguistic input in the indirect form (Bowerman, 1982). The absence of semantically sensible constructions in the surrounding language serves as the evidence of their ungrammaticality.

The Construction Grammar proposal advocated by Goldberg (1995, 2002, 2003, 2006, 2009) is a prototypical example of this research trend. It relies on the theoretical idea of constructions: conventionalized pairings of forms and meanings, including such linguistic units as morphemes, words, as well as sentence constructions. One of the central assumptions of
Construction Grammar is that verbs generally underdetermine the argument structure of a sentence. For example, even though the sentence *She sneezed the foam off the cappuccino* can be understood as involving caused motion of the substance, it is quite unlikely that such interpretation, and, correspondingly, the configuration of the arguments comes from the unique sense of the verb *to sneeze*. Rather, it is argued that the caused-motion reading is directly encoded in the clause-level constructions termed *argument structure constructions*. The argument structure constructions are inductively learned form-meaning correspondences expressing scenes essential to human experience or experientially grounded gestalts such as when someone did something to someone, something moved, someone experienced something, etc. Therefore, verbs and argument structure constructions are interrelated but independent constructional units.

The argument alternation of the locative verbs does not depend on the transformation of one semantic structure into another, but merely indicates use of the verbs in two different argument structure constructions. The content locative frames, such as *spray water onto the plant*, or *load the hay onto the truck* are instances of the locative verbs used in the caused-motion construction. The container locative frames, such as *spray the plant with water*, or *load the truck with hay* are the cases in which the verbs *to spray* and *to load* are used in the causative construction + *with* construction (for further details, see Goldberg, 2002, p. 344).

To explain how language learners come to know which verb can be used in which construction, Construction Grammar theorists advance a usage-based learning account. Goldberg (1995) maintains that a *statistical preemption* is playing an important role in avoiding ungrammatical structures (also, see Braine & Brooks, 1995; Brooks et al., 1999; Theakston, 2004). When a language learner repeatedly hears a verb in a given construction (e.g., content-
locative pour water into the cup), but not in its semantically plausible alternative (e.g., container-locative, *pour the cup with water), the learner infers that the latter is ungrammatical. Thus, the overgeneralizations are preempted or blocked by a consistent use of the verb in one construction when the alternative is expected.

Recent work within the Construction Grammar framework spells out the cognitive mechanism of statistical preemption. It is linked to our capacities to anticipate the upcoming information in the unfolding linguistic stream. For instance, when listening to such sentences as The day was breezy so the boy went outside to fly a … English speakers tend to routinely predict the upcoming word (kite) completing the sentence prior to hearing its final word (for further discussion, see Fine, Jaeger, Farmer, & Qian, 2013; Kamide, 2008; Kuperberg & Jaeger, 2016; Kutas, DeLong, & Smith, 2011). In turn, anticipatory language processing implies at least some degree of indeterminacy: At any given moment in processing there may be multiple plausible continuations of a sentence competing for inclusion into a linguistic structure. When one of the competing candidates wins the competition, the losing representations become inhibited and consequently less accessible (for further details, see Goldberg, Casenhiser, & Sethuraman, 2005; Johnson et al., 2013). The error-driven learning mechanism is necessarily probabilistic. It relies on multiple encounters of the target structures and depends on the distributional properties of the constructions in the linguistic input. Therefore, elimination of the locative ungrammaticalities is the product of the failed expectations that are gradually repressed in the course of linguistic processing.

**Integrating semantics and statistics: A hybrid account**

The two alternative solutions of the learning paradox represent theoretically and ideologically distinct perspectives onto the nature of linguistic development. Yet, many aspects
of the proposals are fully compatible. Both accounts maintain that learning the partially productive generalizations proceeds over an extended developmental period and requires multiple encounters of the target items. Similarly, both proposals agree on a significant role of semantics in determining the ways locative arguments are realized linguistically. The need for both statistical and semantic mechanisms is also apparent from the findings of empirical investigations examining acquisition of argument structure (Ambridge, Pine, & Rowland, 2011; Ambridge et al., 2012; Ambridge, Pine, Rowland, Jones, & Clark, 2009; Bidgood et al., 2014; Gropen et al., 1991a, 1991b; Perfors, Tenenbaum, & Wonnacott, 2010; Twomey et al., 2014; Wonnacott, 2011; Wonnacott, Newport, & Tanenhaus, 2008).

Ambridge et al. (2012) advanced a theoretical proposal integrating semantic and distributional learning mechanisms (also, see Ambridge, 2013; Ambridge et al., 2011, 2009; Ambridge, Pine, Rowland, & Young, 2008). The hybrid account is primarily constructional and views development of the partially productive argument structures in terms of the relationship between verbs and argument structure constructions. The statistical aspects of this relationship are represented as associations between a locative verb and argument structure constructions. Any given verb can appear in various constructional contexts. Correspondingly, hearing a verb activates multiple plausible candidates which complete for the right to accommodate the lexical item. The more often a language speaker encounters a particular verb in a particular construction the stronger is the inference that the unattested uses of the verb are not permitted.

The semantic mechanism is captured as a representational compatibility or semantic fit of verbs and argument structure constructions. On this account, the caused-motion and causative + with constructions can be roughly equated with Pinker’s broad-range semantic categories that represent alternative construal of the locative events. The caused-motion constructions capture
the manner-of-motion aspects of the locative events. Correspondingly, the causative + with construction embodies the state-change perspective. Similarly, the individual verbs or clusters of semantically related verbs are attributed more refined semantic properties reminiscent of Pinker’s (1989) narrow-range classification. As a result, acquisition of the locative alternation patterns can be thought of as learning fine-grained semantic properties of the locative verbs determining the fit between the verbs and argument structure constructions.

**Locative constructions in L2 acquisition**

The semi-regularity of the locative constructions poses the same sort of challenges for L2 acquirers. Provided with inherently limited linguistic evidence, L2 learners are expected to form productive argument realization patterns, which, at the same time, must be restricted to an apparently arbitrary subset of locative verbs. In contrast to the diversity and wealth of L1 investigations, L2 research does not offer any coherent view of how L2 learners might deal with the semi-regularities. In fact, only a handful of studies investigated L2 acquisition of the partially productive patterns, and, most importantly, advanced largely incompatible views on the L2 development of locatives. Juffs (1996) advocated a two-stage lexical parameter account according to which, initially, L2 learners transfer the syntax-semantic correspondences of their native language, but, subsequently, can switch the lexical parameter to the target language settings. In contrast, Bley-Vroman and Joo (2001) and Joo (2003) argued that adult L2 learners are in principle incapable of learning the fine-grained semantic properties of locative verbs (i.e., Pinker’s narrow-range rules) and tend to interpret as alternating even those verbs that do not alternate in their native language. Finally, according to Ambridge and Brandt’s (2013) account, L2 learners do attain a native-like command of locative constructions, but, unlike native speakers, they rely more on the semantic constraints and are less dependent on the statistical
mechanism advocated by the hybrid account. In what follows, I first review these studies in
detail, and then critically examine each investigation. I identify theoretical and methodological
sources of the conflicting claims.

**Juffs (1996)**

Juffs (1996) conducted the first study on how L2 speakers deal with the task of learning the
argument structure of locative verbs. The project focused on cross-linguistic influences on L2
acquisition of locatives. It first examined lexicalization patterns of English and Chinese verbs
revealing systematic differences in what meaning elements become expressed via verbs in each
language. Specifically, Juffs noted that morphologically simple Chinese verbs generally do not
encode caused changes. In contrast to English verbs *to disappoint, to melt, and to cover*, their
putative Chinese translations (*shiwang, ronghua*, and *gai*, respectively) tend to lack the ability to
express the caused change meanings and require additional periphrastic grammatical elements to
form acceptable structures (also, see Kim et al., 1990). The sentences in (4) illustrate the cross-
linguistic differences.

(4)


b. The sun melted the ice. ?Taiyang rong(hua) le xue.
Sun melt ASP snow.

c. John covered the bed with a blanket. ?Zhang San yond tanzi gai le chuang.
Zhang San use blanket cover ASP bed.

Based on this observation, Juffs (1996) proposed a lexical parameter (i.e., +/- caused
change) capturing the incorporation of caused changes by locative and psychological verbs
cross-linguistically. On this account, English and Chinese match in their lexicalization and concomitant argument realization of content-only (e.g., *to pour, to spill*) as well as alternating locative verbs (e.g., *to spread, to load*). At the same time, in contrast to English, Chinese generally lacks container-only verbs (e.g., *to cover, to fill*) or treats them as alternating. As a result, Chinese speaking L2 learners of English should have few difficulties in acquiring either content-only or alternating locative verbs but are expected to show negative transfer effects in learning the container non-alternating locatives. Subsequently, positive L2 input should switch the L1 parameter setting resulting in target-like intuitions about container-only locative/constructions.

To test these predictions, Juffs (1996) conducted an experimental study. One hundred and twenty native speakers of Chinese learning English as a foreign language (of low, intermediate, high, and advanced proficiency levels) and a comparison group of English native speakers completed two experimental tasks. In an elicited-production task, the participants were asked to describe pictures depicting locative events and caused psychological state changes. In the second task, participants judged grammaticality of English sentences including 12 locative verbs in the content-locative and container-locative frames as well as the change-of-state psych verbs in causal and periphrastic constructions. The analyses of experimental results involved one-way ANOVAs comparing the experimental groups of different proficiency levels conducted separately for each category of the target items. Overall, the findings of the production task were interpreted as confirming Juffs’s (1996) hypotheses. As expected, all experimental groups demonstrated statistically comparable performance in describing content-only locative events. At the same time, in comparison to the L1 group, lower proficiency L2 learners (low and intermediate) produced significantly fewer causative psych verbs and were more likely to make
errors describing container-only locative events. The higher proficiency L2 groups did not differ statistically from the L1 participants across any test items. The finding of the acceptability judgement task revealed somewhat different performance patterns. Contrary to the stated predictions, L2 participants in all groups did not differ significantly from English native speakers in their judgments of the container-only verbs in the container-oriented frames (e.g., *Pat covered the bed with a blanket). Neither did they differ statistically from the L1 group in their acceptance rates of the content-only locative verbs in both sentence frames: They rejected the ungrammatical structures (e.g., *Pat poured the glass with water) and judged as acceptable the grammatical content-only sentences (e.g., Pat poured water into the glass). The only difference in L2 speakers’ acceptability judgments were found for the ungrammatical container-only sentences (e.g., *Pat covered the blanket onto the bed). Interestingly, the low and high L2 groups, but not intermediate and advanced groups, differed statistically from the L1 participants in their evaluation of these test items. Despite the challenges in interpreting the experimental results, Juffs (1996) maintained that L2 acquisition of the locative constructions can be best understood in terms of cross-linguistic transfer. Chinese speakers learning L2 English import their L1 lexical parameter settings but are capable of resetting them given sufficient L2 input.

**Bley-Vroman and Joo (2001)**

Bley-Vroman and Joo (2001) conducted an experimental study with a group of Korean native speakers learning English. Fifty-nine advanced L2 learners participated in a forced-choice picture-description task. The study focused on a characteristic property of the container-locative constructions known as the holism effect: the tendency to interpret a container argument as being completely affected by a locative event (e.g., The gardener sprayed the tree with pesticides) (for further details, see Jackendoff, 1996; Pinker, 2009; Tenny, 1994). The experimental task
involved a presentation of a locative verb in either a content-locative or a container-locative frame, along with two story strips illustrating an event. The two illustrations differed in whether the container argument was fully or partially affected. The task was to indicate which picture best represented the meaning of the sentence, with neither as a valid option for ungrammatical structures. The authors reasoned that L2 learners would consistently select a fully affected picture as an illustration of a container-locative sentence and link a partially affected image with a content-locative frame, if they knew the broad-range rules. They also hypothesized that the choice of the neither option for the ungrammatical sentences would indicate the acquisition of the narrow-range rules. The repeated measures ANOVAs, conducted separately for each group of participants, compared the means of selecting the fully affected images for three categories of locatives (content-only, container-only and alternating) used in either content-oriented or container-oriented sentence frame. The outcomes of the ANOVA for the L1 group showed statistically significant main effects of verb category (content-only vs. container-only vs. alternating) and sentence frame (content- vs. container-oriented), as well as a statistically significant interaction between these predictors: The frequency of selecting the completely affected images depended on the category of the locatives and differed between the target verbs in the content and container-oriented sentence frames. Importantly, the difference between content- and container-oriented frames depended on the category of the locatives. In contrast, the ANOVA for L2 participants revealed a significant main effect of the sentence frame and statistically significant interaction between verb category and sentence frame factors but did not show the main effect of the verbal category factor. The advanced L2 learners of English selected the completely affected images more frequently when presented with the locative sentences in the container-oriented sentence frame. This preference for the fully affected images differed for
the content-only, container-only, and alternating locatives. At the same time, the means did not differ among different categories of locative verbs (when the sentence frames were not taken into account). The authors interpreted the experimental results as indicating that the advanced L2 learners of English had knowledge of the broad-range rules, but failed to reject ungrammatical sentences, therefore, demonstrating a lack of the narrow-range rules that constrain the alternation patterns of the locatives. The Korean learners of English tended to interpret and relate to the pictures the non-alternating structures such as *John poured the plant with water even though the content-locative translation equivalents of English pour, spill, glue, and nail, are also non-alternating in Korean.

Joo (2003)

Joo (2003) published a follow up article reporting results of a second experimental task with the sample presented in Bley-Vroman and Joo (2001). The participants completed a forced-choice sentence-selection task, which differed from the previous experiment in that two sentences were presented with a single picture strip. The L2 learners were to select a sentence that best described an illustrated event. Similarly to Bley-Vroman and Joo (2001), Joo (2003) used repeated measures ANOVAs conducted separately for L1 and L2 participants. The analyses compared the mean number of container-oriented sentences chosen as the descriptions for either completely affected or partially affected images depicting content-only, container-only or alternating locative events. Curiously, the statistical outcomes demonstrated the same effects for both experimental groups: (a) participants were more likely to choose the container-oriented constructions depicting completely affected events, (b) their overall choice of the container-oriented sentences differed across categories of the locative verbs, and (c) the preferences for the container-oriented constructions as descriptions of the fully affected images depended on the
category of the locatives. Despite such statistical outcomes, Joo (2003) maintained that the findings of this study generally aligned with the results of the previous experiment concluding that advanced Korean learners of English did not display knowledge of the narrow-range rules.

In sum, Bley-Vroman and Joo (2001) and Joo (2003) maintained that the L2 acquisition of the locative constructions might be categorically different from L1 learning. Therefore, they outlined a rather pessimistic scenario of acquiring the partially productive patterns in L2 holding that “[…] principled knowledge of narrow classes is difficult or impossible to attain, although learners may be able to associate (in an “unprincipled” way) given verbs with given constructions, owing to input exposure or even to explicit presentation” (p. 216).

Ambridge and Brandt (2013)

Finally, Ambridge and Brandt (2013) conducted a grammaticality judgment study with 30 native German speakers who were advanced L2 learners of English. The project tested the hybrid account of learning the partially productive locative constructions. The study used mixed-effect modeling to estimate effects of the semantic predictors\(^1\) (broad and narrow-range classes) as well as statistical properties of the target verbs on participants’ acceptability judgments. The participants’ task was to rate on a 7-point Likert scale the acceptability of 60 English locative verbs used in content-oriented and container-oriented sentence frames. The dependent variable was a difference score calculated by subtracting a rating for a container-oriented sentence (e.g., *Pat poured the glass with water*) from its content-oriented counterpart (e.g., *Pat poured water*).

\(^1\) The semantic predictors consisted of seven composite measures computed in a previous normative experiment (Ambridge et al., 2012). They indicated the extent to which each locative verb exhibits broad-range (“manner”, “end-state”) and narrow-range (“splattering”, “joining”, “stacking”, “gluing”, and “smearing”) semantic properties.
into the glass). The experimental findings provided support for the authors’ account. Both semantic and statistical characteristics of the target verbs were found to independently predict the participants’ acceptability ratings. However, in contrast to the native speakers of English examined in a previous investigation (Ambridge et al., 2012), the advanced L2 English learners displayed somewhat different results. The semantic “end-state” variable turned out to be only marginally significant ($p = .06$) as a predictor of the L2 participants’ judgements suggesting relative insensitivity to the ungrammatical uses of container non-alternating locatives (e.g., Pat filled water into the glass). Ambridge & Brandt (2013) interpreted this finding as evidence of L1 transfer. Also, unlike the L1 English speakers examined in Ambridge et al. (2012), the L2 learners did not demonstrate the frequency effects predicted by the Construction Grammar proposal. The verbal frequencies were not associated with a reduced acceptability of the unattested (i.e., ungrammatical) locative constructions as suggested by the statistical preemption proposal. Rather, the grammatical test items of higher frequencies were rated as more acceptable by the L2 participants. Based on these findings, the authors came to the following conclusion: Even though the L2 participants demonstrated the near-native-like performance on the experimental task, their greater tolerance of the ungrammatical constructions (especially container-only ill-formed structures) as well as the lack of the predicted statistical effects suggested that “the L2 speakers appear to have arrived at their near-native-like command of the English locative construction via a somewhat different route” (p. 255). Rather than relying on the gradual depreciation of the unattested/ungrammatical constructions as the result of witnessing the well-formed items, L2 learners depend more on the presumably underdetermined and affected by their native language verb and construction semantics.
Critical review

The empirical investigations on the learning-argument structure of locative verbs in a L2 arrived at conflicting conclusions. Table 1 summarizes the L2 studies. In the following critical analysis, I identify theoretical and methodological inconsistencies that could potentially account for the mutually exclusive claims.

Table 1. Summary of empirical studies investigating locative verb constructions in L2

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Experimental Task</th>
<th>Analysis</th>
<th>Learnability of Locatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juffs (1996)</td>
<td>L1 English and low, intermediate, high and advanced L2 English learners</td>
<td>AJT</td>
<td>between-group</td>
<td>Yes</td>
</tr>
<tr>
<td>Bley-Vroman and Joo (2001)</td>
<td>L1 English and advanced L2 learners</td>
<td>sentence-picture matching</td>
<td>within-group</td>
<td>No</td>
</tr>
<tr>
<td>Joo (2003)</td>
<td>L1 English and advanced L2 learners</td>
<td>picture-sentence matching</td>
<td>within-group</td>
<td>No</td>
</tr>
<tr>
<td>Ambridge and Brandt (2013)</td>
<td>Advanced L2 learners</td>
<td>AJT</td>
<td>within-group</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*Note, AJT = acceptability judgment task*

Juffs (1996) attempted to explain the complex argument structure phenomena in terms of a quite elegant theoretical device. The parameter links superficially diverse verbs to the abstract properties of the lexical items and explains lexicalization patterns within and across languages. Such an attractive undertaking, however, is founded on internally inconsistent assumptions and lacks unequivocal empirical support.
Juffs (1996) remained rather vague about the exact status of the Chinese translation equivalents of English container-only locative verbs (e.g., Chinese locative *gai* ≈ *to cover*). On some occasions, the author refers to these lexical items as not capable of accommodating caused changes. On other occasions, these verbs are treated as either content-only or alternating locatives. Provided that these verbs readily accommodate Content (but not Container) arguments as their direct objects, we can treat them as content-only locative verbs. This raises at least two further problems. First, it is not clear on what basis the Chinese equivalents of English *to fill* and *to cover* (being content-only verbs in Chinese) relate to the psych verbs and why other Chinese content-only verbs that have identical syntax-semantics characteristics cross-linguistically (e.g., *to pour*) escape the relationship. Correspondingly, it is puzzling how the proposed parameter is expected to work in language acquisition acting only on the subset of the Chinese content-only verbs (i.e., of *cover* and *fill* kind) but ignoring the content-only *pour*-like verbs. The same argumentation applies if we treat Chinese *cover/fill*-like verbs as alternating. It is unclear how these and only these hypothetically alternating lexical items relate to the Chinese verbs incapable of encoding caused state changes, and what prevents the lexical parameter from affecting the rest of the alternating verbs behaving similarly in English and Chinese.

Even more problematic for Juffs’s (1996) proposal are the experimental results: Neither did they corroborate the parametric account, nor they warrant the author’s inferences. In the guided production task, the low and intermediate L2 learners indeed showed statistically different performance from the L1 group. However, these differences were most likely the differences in a degree of exhibiting the target-like performance. For example, describing the container-only locative events (*to decorate, to cover, to stain, and to block*) the L2 learners did show the target-oriented performance producing such sentences as *The man decorated the tree*
with the toys (Low L2, M = 1.00, SD = 1.63; Intermediate L2, M = 0.81, SD = 0.87). They differed from the native speakers (M = 2.03, SD = 0.51, maximum score of 4), but, importantly, did not use their L1 patterns.

Even if readers assume that the reduced production of target sentences is indicative of the L1 transfer, the results of the acceptability judgement task are difficult to interpret. Contrary to the parametric predictions, L2 learners rated both content and container non-alternating grammatical sentences as highly acceptable. The low proficiency L2 group showed the lowest ratings of the container-only structures (e.g., *The man filled the glass with water*) on a seven-point Likert scale ranging from -3 to +3 with “+3” representing a “completely possible” choice (Low L2, M = 2.25, SD = 1.45). Nonetheless, they patterned with the L1 speakers (L1, M = 2.95, SD = 1.15). Similar results were found for Low L2 learners (M = 2.06, SD = 2.79) and native English speakers (M = 2.79, SD = 0.40) on their ratings of the content-only items (e.g., *The man poured water into the glass*). Therefore, even the low L2 group showed target-like performance. They judged English content and container non-alternating sentences as fully acceptable. The only statistically significant results that could be putatively interpreted as evidence in support of the lexical parameter were reported for the participants’ ratings of ungrammatical container-only constructions. The low proficiency L2 learners were less likely to reject such ill-formed sentences as *The man filled water into the glass* when compared to the L1 group (Low L2, M = -0.25, SD = 1.72 vs. L1 group, M = -1.77, SD = 0.86). This difference, however, was rather small, or perhaps even meaningless, as both L1 and L2 speakers’ mean scores were within the negative portion of the rating scale. Therefore, rather than supporting the parametric proposal, the experimental results point in the opposite direction: Even beginning L2 learners are capable of producing and accepting sentences containing locative verbs that incorporate caused changes.
They also reject the ill-formed container non-alternating constructions but do demonstrate greater tolerance of the violations.

Bley-Vroman and Joo (2001) and Joo (2003) claimed that adult L2 learners, unlike children learning their native language, are incapable of acquiring the partially productive locative structures. Schwartz, Dekydtspotter, and Sprouse (2003) offered a full-fledged critique of this proposal pointing to a number of theoretical and methodological weaknesses as well as errors in interpreting the statistical results of the empirical studies. Taken together, these shortcomings largely undermine Bley-Vroman and Joo’s (2001) and Joo’s (2003) proposal.

The essential assumptions of Bley-Vroman and Joo (2001) and Joo (2003) investigations were that (a) the sensitivity to the holism effect indicates participants’ knowledge of the broad-range rules, and (b) accuracy in the sentence-picture matching tasks (i.e., selecting “neither” option for ungrammatical items) reflect their knowledge of the narrow-range rules. Schwartz et al. (2003) argued that neither of these assumptions withstands a closer examination. The holism effect is a primarily pragmatic inference and, in many cases, depends on the context (for detail, see Jackendoff, 1996). The contextual dependency of the holistic interpretations weakens the hypothesized relationship between the knowledge of the broad-range rules and selection of the visual images. More importantly, it calls attention to the convoluted nature of the experimental tasks blending together acceptability and interpretability of the test items. Schwartz et al. (2003) point out that acceptability and interpretability are not fully synonymous concepts: language speakers can interpret ungrammatical structures and have difficulties in interpreting fully grammatical ones.

The problems with the experimental task become apparent if we consider the experimental outcomes. In the forced-choice picture-description task, reported in Bley-Vroman
and Joo (2001), participants related a single test sentence to one of the two visual images of locative events. The L2 learners ($M = 2.83$, $SD = 1.24$) but not native English speakers ($M = 0.59$, $SD = 0.94$) tended to interpret such test sentences as *John poured the glass with water* as signifying the fully affected locative scenes. This finding is the main source of the lack of statistically significant effect of the verb class found by Bley-Vroman and Joo (2001). The authors interpreted the non-significant result as evidence that L2 learners do not distinguish between alternating and non-alternating locatives, and, thus, do not have knowledge of the narrow-range rules. Curiously, in the second task reported in Joo (2003), where participants were given two sentences and could choose one to match with a picture, the statistical analyses showed similar results for L2 and L1 groups: Both groups showed statistically significant effect of the verb class demonstrating differentiation between alternating and content- and container non-alternating verbs. Even though the L2 group did tend to perceive as if grammatical the ungrammatical content-only structures ($M = 2.47$, $SD = 1.37$) to a greater extent than L1 speakers ($M = 0.41$, $SD = 0.71$), the statistical tests conducted separately for each group showed that the participants’ performance in each group depended on the class of the locative verbs.

Ambridge and Brandt’s (2013) investigation is based on a more flexible theoretical foundation which maintains that both semantic and statistical factors play an important role in the acquisition. Correspondingly, unlike Juffs (1996) or Bley-Vroman and Joo (2001) and Joo (2003), this study did not commit to any strong predictions and explored the L2 participants’ ratings of the locative constructions in terms of the hybrid account. Based on the experimental results, the authors did come up with quite strong characterizations of L2 learning and suggested that the L1 might influence the participants’ judgments in L2, and that L2 learners are likely to reach native-like knowledge via a different developmental path. These inferences, however, do
not necessarily follow from the empirical results. The authors interpreted the marginally significant effect of the “end-state” semantic predictor as evidence of presumable L1 transfer. Although plausible, such interpretation is not obvious: it is not clear how the marginally significant predictor relates to the composite outcome variable (i.e., difference score) so that signaling the L2 transfer effect. Curiously, in a previous study of L1 English speakers, the participants showed the reversed statistical pattern. For this sample, the “state-change” semantic predictor was significant, but the other broad-range semantic variable (i.e., “manner-of-motion”) failed to reach the significance level \(p = .08\). This finding, however, was not interpreted as demonstrating any deficiencies in the English speakers’ competence. Rather, it was taken as merely indicating the relevance of these broad-range semantic properties of the locative verbs to the participants’ rating performance. Therefore, it remains unclear how similar statistical outcomes can be used to attribute opposite characteristics to the L1 and L2 populations.

Similarly challenging is to attribute L2 learners a greater reliance on the semantic learning mechanism based on the effects of the statistical learning predictors which, nevertheless, did not corroborate the theoretical account. The study found consistent frequency effects for the well-formed locative constructions: The L2 learners tended to rate as more acceptable the test items which they witnessed more frequently. The authors acknowledged this observation but did not attempt either to revise their theoretical assumptions or to reason through why these findings are not relevant for the characterization of L2 learning. Rather, they focused on the alternative semantic learning route and largely disregarded the statistical learning effects.

**Methodological sources of the conflicting accounts**

An additional source of the incongruent conclusions of the empirical studies lies in the way the investigators defined the dependent measures and used them in the statistical analyses.
Juffs (1996) measured participants’ perception of the locative constructions on the 7-point Likert scale and compared the L2 and L1 experimental groups separately for each category of the locatives used in the content- or container-oriented sentence frames. A total of six one-way ANOVAs were used to assess the participants’ knowledge of the target locative structures. In Bley-Vroman and Joo (2001) and Joo (2003), the researchers used the interpretative scores, which measured the participants’ sensitivity to the complex relationship between locative sentences and their visual representations. In contrast to Juffs (1996), the statistical analyses were conducted separately for each experimental group. Finally, Ambridge and Brandt’s (2013) project relied on a difference score which quantified the degree to which the participants preferred or disfavored structural realizations of the locative verbs in the content- and container-oriented sentences. The authors’ statistical analyses estimated the effects of the semantic and statistical predictors on the dependent measure and were conducted within a single group of participants.

Each depended measure has its own intuitive appeal and reflects unique aspects of the participants’ performance. However, all of them have many potential pitfalls. Juffs’s (1996) comparisons of the L1 and L2 participants allow incorporating gradience of the acceptability judgments, but, at the same time, require careful classification of the observed outcomes: The between-group differences can lay anywhere on the scale indicating either true divergences or differences in the degree to which the participants accept or reject the test items. Bley-Vroman and Joo (2001) and Joo’s (2003) depended measures condense interpretability and acceptability into a single metric, but, simultaneously, conceal complex interaction between these overlapping but non-tautological aspects of perceiving linguistic structures. In turn, the within-group statistical analyses make it difficult to draw definitive inferences about the between-group
differences and/or similarities leaving considerable room for ad hoc speculations. The difference
scores used by Ambridge and Brandt (2013) allow collapsing participants’ ratings of locative
verbs in two alternative frames and thus enable researchers to examine the effects of multiple
predictors on the outcome variable in a unified manner. However, similarly to Juffs’s (1996)
comparisons, the difference scores leave underspecified where these differences come from. As a
result, it is difficult to draw clear and unambiguous conclusions about specific effects of the
predictors on the participants’ rating performance.

The research on the acquisition of argument structure in L2 offers a rich source of
empirical data that can be examined in different ways. In a simplest case, participants’ ratings
can be evaluated in absolute terms, regardless of their judgements of the target items in
alternative syntactic constructions or performance of L1 or L2 speakers in the comparison
groups. Therefore, researchers can examine the likelihood of accepting a particular kind of test
items in a particular construction (e.g., *John filled water into the glass). Alternatively, the
scores can be compared across groups of participants of different L2 proficiency levels or with
respect to the experimental performance of native speakers. Such between-group or benchmark
comparisons, however, should be interpreted with care and cross-validated by additional
examinations of the outcomes in absolute terms. Also, the acceptability judgments can be
examined in preferential terms: whether or not participants within a single group demonstrate
any preferences for one syntactic realization of the test items over the other. For example, the
analysis can compare ratings of such constructions as John filled the glass with water vs.*John
filled water into the glass, in such way examining whether a particular group of participants
distinguishes between alternative realizations of the test items. Similar to the benchmark
comparisons, the preferential examinations require careful handling especially when drawing
conclusions about categorical properties of the participants’ ratings evident only when their judgments are assessed in absolute terms.

In sum, the methodological choices made by the investigators played an important role in advancing the irreconcilable proposals. In many cases, the studies’ conclusions were considerably biased by what the researchers chose to observe and how these observations were handled in the statistical examinations.

**Conclusion**

A closer look at the L2 investigations suggests that neither of the simpler solutions of the learning paradox advocated by Juffs (1996) and Bley-Vroman and Joo (2001) and Joo (2003) can account for the observed experimental findings. It is unlikely that L2 learners acquire specifics of the locative verb constructions by transferring their native language syntax-semantics correspondences and subsequently switching the lexical parameter to the target setting. Nor is it the case that they are imperceptive of the fine-grained distinctions expressed by different categories of the locative verbs. Rather, L2 learners in many respects resemble young language acquirers: The main locus of non-convergence with the target language standards appears to be in a greater tolerance to the grammatical violations. Unlike native speakers, L2 learners more readily interpret and accept the ungrammatical forms of the non-alternating locatives. The greater tolerance for the ungrammaticalities can be a product of L1 influence or reflect true maturational differences between young L1 learners and adults acquiring their L2. Consequently, mush weaker versions of Juffs’s (1996) and Bley-Vroman and Joo’s (2001) and Joo’s (2003) accounts can be considered in fully compatible and non-conflicting terms as proposals characterizing L2 development of the locative argument structure.
The reconciliatory perspective eliminates the apparent contradictions but at the same time pushes L2 research on the learning paradox back to square one: It introduces a range of uncertainties about processes and outcomes of learning the “unlearnable” locative constructions. Do L2 learners fully recover from the overgeneral argument realization patterns? If yes, how they do so and what factors affect L2 development in either positive or negative ways? Are L2 learners capable of relying on statistical properties of the locative constructions or do they solely depend on learning the semantics of locative verbs? If L2 learners’ linguistic development does differ from L1 acquisition, how do we account for the different developmental routes? None of these questions has received any definite answers or can be effectively assessed by the available empirical evidence.

The inconsistencies of the L2 empirical studies are rooted in their theoretical underpinnings. Both Juffs’s (1996) as well as Bley-Vroman and Joo (2001) and Joo’s (2003) accounts included hypotheses about the states of L2 learners’ grammar (i.e., initial/final) without any means to address how these states come or fail to come about. In Gregg’s (2008) terms, these researchers advanced the property theoretical proposals leaving unspecified the transitional characteristics of L2 development (for further details, see Gregg, 2005; Jordan, 2004). Similarly, Ambridge and Brandt’s (2013) account, in its current formulation, has no way to account for why L2 learners’ performance diverges from L1 speakers’. Everything else being equal, the hybrid account should predict a full convergence in the performance of L2 and L1 speakers rather than deficiencies of the statistical and semantic learning mechanisms in L2 development.

The gaps in the theoretical formulations, in effect, licensed the apparent selectivity in the statistical analyses, and, most importantly, biased interpretations of the empirical findings. The
studies tended to focus on those aspects of the empirical results that fit their theoretical orientation, and generally disregarded the inconvenient discrepancies. Therefore, any theoretical account of learning locative constructions in L2 should be able to accommodate not only hypotheses about L2 learners’ grammar(s), but also specify learning processes by which such grammar(s) might be acquired. In other words, a coherent theory must spell out how language learners’ linguistic experiences relate to their grammars so that resulting or failing to result in the developmental changes.

**A processing perspective on learning locative constructions**

The foundational assumption of the current proposal is that any solution of the learning paradox must account for how language learners’ real-time linguistic experiences relate to the complex, semi-regular linguistic generalizations. In some sense, such assumption harks back to the question posed by Baker (1979) in his original postulation of the learning paradox: “What is the functional relation that exists between an arbitrary being’s early experience (his “primary linguistic data”) and his resulting adult intuitions?” (p. 533). The crucial difference between otherwise identical concerns is that the current proposal calls for an online processing interpretation of the term “linguistic experiences”.

The traditional discourse views a set of such constructs as *environmental language*, *linguistic evidence*, *primary linguistic data*, or *linguistic input* as primarily external entities that contain information about linguistic structures. Consequently, many debates center around such issues as whether these linguistic objects are too impoverished to determine the complexity of our linguistic knowledge, or, alternatively, contain a range of statistical, distributional, or other kinds of cues enabling the course and outcomes of language acquisition. In contrast, viewed from an information processing perspective, the abstract objects cannot sustain their integrity. The
environmental contribution is limited to physical signals or *stimulus arrays* (visual, auditory, etc.) which are inherently void of familiar linguistic constructs: There are no phonemes, syllables, words, grammatical features, or conceptual structures in the continuous vibrations of the air or currents of light particles (for a related discussion, see Jackendoff, 1989). All these items are internal mental structures constructed in the mind for the purpose of linguistic communication by re-representing stimulus arrays across multiple processing stages or levels (Carroll, 1999, 2001, 2017a, 2017b; Sharwood Smith & Truscott, 2014; Truscott, 2014).

Correspondingly, our linguistic experiences are nothing more than real-time construction of the mental linguistic structures. Even though we can habitually speak of morphemes, words or sentences heard by a language learner, they are convenient abstractions from mental states and processes underlying language comprehension.

Viewing linguistic experiences from a real-time processing perspective requires a further qualification of these phenomena. We perceive spoken sentences as a succession of linguistic units. In other words, the mental structures are built *incrementally* (Rayner & Clifton, 2009) in order to update to integrate the upcoming information as rapidly as possible. The bottom-up moment-by-moment processing is *interactive* (Chang et al., 2013; Pickering & Garrod, 2007, 2013), as it might be influenced by representations on the higher levels of processing. For example, semantic or pragmatic factors might affect the course of syntactic or phonological processing. Furthermore, real-time comprehension is highly *competitive*: on each processing level, there are multiple candidates competing for integration in a representation of a sentence. Additionally, to deal with the rapidly unfolding signal in the most efficient way, online processing is not limited to integration of the upcoming signal but also involves *prediction* or *anticipation* of what comes next in the speech stream (Kamide, 2008; Kutas et al., 2011). Finally,
the processing events tend to facilitate subsequent encounters of the same or similar linguistic constructions. In other words, real-time processing is *adaptive* in its nature (Fine et al., 2013).

Therefore, the current approach construes linguistic experiences as information processing events which display complex dynamic characteristics of an effective and efficient communication system. Such system is necessarily reciprocal: Our linguistic experiences are determined and limited by the mental mechanisms and cognitive capacities enabling the processing events. At the same time, the only way to acquire and develop the linguistic structures is through experiencing (i.e., processing of) meaningful samples of a language. Consequently, the questions of learning the locative constructions boil down to a specification of the cognitive machinery underlying the linguistic processing as well as ways the processing events can lead to the growth of the linguistic structures. The following section presents a theoretical framework capable of articulating development of the partially regular patterns in processing terms.

**Modular Online Growth and Use of Language (MOGUL)**

Modular Online Growth and Use of Language (MOGUL) is a theoretical framework that enables a unified examination of linguistic structure, processing and development (Sharwood Smith, 2017; Sharwood Smith & Truscott, 2014; Truscott, 2014; Truscott & Sharwood Smith, 2004). Initially inspired by Ray Jackendoff’s Parallel Architecture model (1983, 1987a, 1990b, 1997, 2003), MOGUL views the mind as a network of specialized cognitive systems (i.e., *modules*), each dedicated to handling a specific kind of information. Any cognitive task, including comprehension and production of a language, is understood as the functioning of a coalition of the domain-specific “expert systems”.

36
Linguistic knowledge is the product of semi-independent “expert systems” dealing with linguistic sound (phonological module), syntactic structures (syntactic module), and semantic/conceptual linguistic structures (conceptual module) (Figure 1). Each module consists of a memory store and an integrative processor. The memory stores are repositories of the module-specific elements and their combinations. The phonological store contains phonological primitives as well as other structured representations of linguistic sounds (viz, phonemes, syllables, morae, phonological words and phrases). The syntactic and conceptual stores hold syntactic and conceptual structures, respectively. The function of the integrative processors is to build complex structures from the elements in their respective stores. For instance, the phonological processor composes phonological structures independently from the syntactic and conceptual processors. Correspondingly, the syntactic and conceptual processors deal with the construction of the representations in their respective stores. Each module operates using its own code and procedures independently from the adjacent units. Such design ensures effectiveness and efficiency in information processing and development. Given a set of innately specified primitives (a representational starter kit) and a set of innate and invariant compositional
algorithms, each module constructs the mental structures within a relatively short developmental period and can efficiently deal with a prescribed set of tasks in real-time processing.

The representations of the independently functioning modules are interrelated via the *interface processors* (double-headed arrows in Figure 1). The main function of the interfaces is to match structures in the adjacent memory stores. A single lexical item or a complex sentence can be thought of as a chain of representations. For example, the word *cat* is the associated bits of the phonological /kæt/, syntactic (+N, singular), and conceptual (ANIMATE, ENTITY, FELINE) structures. Additionally, non-linguistic information (such as products of perceptual “expert systems”) might be a part of the representational chain, including visual, auditory and tactile memories linked to the lexeme *cat*. Therefore, linguistic knowledge is the function of semi-independent “expect systems” and is ultimately realized as representational chains.

The view of language as a collaborative product of the parallel generative modules facilitates reinterpretation of abstract linguistic knowledge in processing terms. The construct of *activation* serves an important role in MOGUL’s theorizing. The activation metaphor refers to the availability of representational elements in the modules’ memory stores: The more active a representation is, the more readily it can be used by a dedicated processor. The highly-activated structures that pass an activation threshold (i.e., the items in the *working memory* state) (e.g., see Cowan, 2012) are the structural elements used by processors to form novel composite representations. Thus, any processing event within a particular module can be seen as a construction of cognitive representations from the activated representational ingredients.

Any language processing event involves building mental structures across a range of modules. The comprehension of a spoken sentence starts with processing of the environmental sensory input by the auditory module (or auditory processing unit). The auditory processing unit
deals with frequency, duration, amplitude, etc. of the acoustic signal and builds mental representations of the environmental sounds (Carroll, 1999). The output of this “expert system” is what we experience as an array of environmental sounds of both linguistic and non-linguistic nature. At each moment of processing, the auditory structures (AS) in working memory raise the activation level of the relevant phonological structures (PS) in the phonological module via the dedicated interface processor. Correspondingly, the interface processors linking PSs and syntactic structures (SS), and SSs with conceptual structures (CS) trigger the structures in the adjacent processing units raising their activation level and thus making them available for the dedicated integrative processors. In other words, the representations activated/built by the dedicated processors within a given expert system, (i.e., their representational outputs), can be thought of as input for the adjacent modules.

This process is necessarily incremental: at each moment of language comprehension the integrative processors work with whatever is available in the module’s working memory. The incrementality of processing also implies that the interface processors activate a substantially larger number of structures, only a handful of which subsequently become a part of the representational chain: Within each module there are many potential candidates competing for inclusion in the final representation. Furthermore, the process of language comprehension can be influenced by top-down processes. For example, visual information (i.e., the product of the visual module) and conceptual structures can affect the bottom-up processing by raising the activation level of the linguistic structures via conceptual/syntactic interface processor. The top-

---

2 The current description of speech perception provides a somewhat simplified view of the auditory processing. A number of factors such as top-down influences, interaction between auditory and visual modalities, as well as a listener’s L1 are likely to affect his/her processing of language.
down processes, given the right conditions, allow to predict the upcoming information and in doing so that constrain the incremental and competitive dynamics of language comprehension. In sum, any instance of linguistic processing involves cascading activation of representational components across multiple domains and incremental construction of mutually constraining cognitive structures.

The key feature of the MOGUL framework is a unification of processing and acquisition. The Acquisition by Processing Theory (APT) is an integral part of MOGUL. It views acquisition and growth of cognitive structures as a byproduct of real-time processing (for summary, see Sharwood Smith, 2010; Truscott & Sharwood Smith, 2004). The APT construes linguistic development in the following terms. The mental structures as well as the links between the representations constructed by the semi-independent “expert systems” are all thought to have a certain level of activation (i.e., mnemonic accessibility) even when they are not in use. This resting level of activation largely determines how easily a particular representation can win competition and be used in a representational chain. The mental structures used in an ongoing processing event acquire a high (current) activation level, but then swiftly fade out of working memory. They gradually descend to a new resting level of activation slightly above the level prior to the processing event. In other words, each use of a mental structure in real-time processing strengthens this representation making it more accessible for the following encounters. This equally applies to the primitive as well as complex representational units including features, words and complex sentences. The complex constructions however would usually have a relatively low level of activation and in most cases, fail to become well-established linguistic units.
Linguistic development in MOGUL terms can be best understood as an adaptive phenomenon: The adaptation is guided and constrained by the nature of the “expert systems” and ultimately reflects language learners’ previous experiences. Such learning in some sense bears characteristics of a Darwinian process. The representations established by the different “expert systems” as well as the links between these structures survive the fast-paced and competitive linguistic dynamics as long as they can serve as building blocks for future processing events. In a nutshell, the growth and development of cognitive representations in MOGUL is seen as the lingering effect of processing.

**Acquisition of locative verbs’ argument structure**

The semi-regularity of the locative verbs’ argument structure can be readily reinterpreted within the MOGUL framework. In essence, the specifics of the locative constructions can be captured as the relationship between the PS⇔SS part of the representational chain and CSs encoding verbal meanings. The main issue in the development of the locative constructions is how language learners arrive at the right semantics.

From a MOGUL perspective, the overgeneralizations as well as retreat from the overly general patterns are the natural byproducts of linguistic processing. By definition, any locative verb denotes a non-stative relationship between Containers and Content caused by an Agent. In turn, any locative event can be construed as an instance of either caused motion or a caused

---

3 Unlike L1 acquirers, L2 learners are more likely to receive a substantial amount of direct negative evidence in the form of either corrective feedback or grammar instruction. I remain agnostic about a role of the instructional interventions in the development of locative constructions. The current empirical evidence is insufficient to draw any conclusions about effectiveness of such treatments. Moreover, it is unclear how L2 practitioners can organize systematic instructions targeting locative argument realization patterns.
change of state. Therefore, the availability of both interpretations is the default option for any locative verb. On the initial encounters of novel locatives, presumably heard in the context of visually presented actions, language learners are faced with an apparent ambiguity given that either Content or Container can be perceived as an affected constituent and thus take the direct object position in the linguistic structure. The instances of overgeneralization are likely to be observed throughout the course of language learning and can be thought of as the signs of developmental optionality (for a processing account of optionality, see Truscott, 2006). Provided a number of contextual influences such as the pragmatics of a particular situation, previous discourse, size of a person’s lexicon, a language learner might overextend the meaning of a locative verb and produce ungrammatical constructions. For example, in a situation when a Content is unintentionally transferred onto a particular surface, a speaker might produce …*spilled the table cloth with …* instead of grammatical …*spilled coffee onto the table cloth.*

The specifics of the occurred event (e.g., ruining the aesthetic qualities of the receptacle), the question *What happened to the table cloth?* asked by an interlocutor, or merely a lack of better lexical means to deliver the intended message might bias a speaker to produce the unconventional construction. Similarly, a developing language user presented with the ungrammatical locative construction is likely to judge it as an acceptable structure merely because both semantic interpretations are initially available for all locatives.

The growth of the semi-regular locative constructions is a matter of fine-tuning the form-meaning correspondences. Such a process necessarily involves both semantic and statistical learning and ultimately instantiates the acquisition-by-processing mechanisms. The semantic development involves refining of the verbal meanings via *cross-situational learning:* the procedure of reducing ambiguity of verbal meanings through witnessing lexical targets across
multiple linguistic and non-linguistic contexts (Pinker, 1989; Smith, Suanda, & Yu, 2014; Yu & Smith, 2007). In the MOGUL’s terms, the contexts in which the semantic development takes place are mental representations constructed across various “expert systems”. These representational ensembles accompany establishment of the target items and collectively make more salient (i.e. available for integration during linguistic processing) conceptual elements distinguishing different subclasses of the locatives. A possible source of contextual influence is a language discourse preceding a target structure. The sentences heard or read before the targets are essentially the PS⇔SS⇔CS chains constructed by the corresponding processing units. The necessary product of such processing is the sets of coherent conceptual structures whose activation levels remain elevated during processing of locative constructions. Jointly with other active conceptual representations, such as a listener’s goals, his/her construal of interlocutors’ goals and intentions, and/or cognitive scripts (stereotyped sequences of events), the CSs of the prior discourse narrow down semantic aspects of upcoming verbs. As such, the “conceptual contexts” facilitate language learners’ inferences about meanings of the target verbs by pre-activating (i.e., making more accessible) semantic elements that are characteristic of specific items or sub-classes of locative verbs.

A more direct role plays sensory-motor representations: a set of mental structures built across dedicated modules dealing with perception and action of an organism in the environment (for example, see Pulvermüller & Fadiga, 2010). Each time a learner hears a locative sentence in the context of an unfolding environmental event, the modality specific representations (e.g., visual, auditory, haptic, etc.) become associated with the conceptual structures built in the course of sentence processing. The associations are the workings of the dedicated interface processors whose job is to match active modality specific representations with the corresponding conceptual
elements. The consistently reenacted associations grow into integral parts of the representational chains, and, as a result, delimit the semantic scope of the linguistic items by increasing mnemonic accessibility of specific aspects of the locatives’ semantics. For instance, distinctive visual patterns of substances’ motion (visual representations) as well as accompanying sounds (auditory structures) become invariable perceptual coordinates of such verbs as to pour, to drip, and to drizzle through repeated processing of the linguistic constructions in the context of perceived locative events. In such way, the invariables “prime” (i.e., raise activation level) the manners of motion as the essential component of the verbal meanings. In contrast, a different set of perceptual structures becomes associated with such verbs as to fill and to cover. Namely, particular changes occurring to receptacles (but not the ways these changes come about) are the common percepts repeatedly reactivated across different uses of these verbs. Correspondingly, the state change of Containers becomes the prominent aspect of the verb’s meanings.

Tightly coupled with the perceptual processing is the functioning of motor structures: mental mechanisms underlying our capacities to plan and execute actions as well as recognize and understand actions performed by others (e.g., see Fischer & Zwaan, 2008; Glenberg & Gallese, 2012). The motor representations primarily contribute to the development of the force-dynamic aspects of the verbal meanings: the semantic dimension specifying a family of causal relations (e.g., forcing, helping, enabling, and hindering) (Deane, 1996). Even the most rudimentary motor actions, such as touching or reaching for an object, involve complex computation of the amount, duration and direction of force exerted by the muscular system to/toward an entity as well as probable outcomes of the force application. In close collaboration with the perceptual systems, these representations afford growth of the fine-grained conceptual composites specifying whether locative actions are accomplished by enabling Content’s motion
(pour, drip, drizzle), forcing contact and motion of Content against Container’s surface (spread, smudge, smear) or by applying force launching ballistic motion of Content in a specified direction (splash, sprinkle, splash). Similarly to the perceptual structures, the motor structures trigger the conceptual development via interface processors matching mental structures between the independently functioning modules.

The conceptual and sensory-motor contexts do not necessarily exhaust all possible sources of the cross-situational learning. The mental mechanisms underlying the affective domain as well as complex interactions among perception, action and affect are likely to drive further refinement of the verbal meanings. Nevertheless, any contextual effects follow the same basic logic. They arise from the continuous coordination of the linguistic processing with the mental structures built by the cluster of relevant “expert systems”.

In sum, the nuanced verbal meanings are the byproduct of language processing. They are established and gradually refined through real-time communicative events. As such, these conceptual representations cumulatively incorporate the semantic components compatible with a range of contextual coordinates and ultimately reflect language learners’ linguistic experiences. This includes both situated, contextualized learning, when L2 learners observe or participate in activities described by locative verbs, as well as learning from exposure to printed material.

Statistical learning mechanisms play a key role in fine-tuning the form-meaning correspondences of locative verbs. In MOGUL, the statistical effects are accounted for in terms of activation of mental representations. Each use of a mental structure in a processing event increases its availability (i.e., resting activation level) for the following processing. The frequently (re)used representations gradually acquire a high resting activation level, and eventually grow into well-established units. Correspondingly, the mental structures constructed
for on a given occasion but not utilized in the future steadily lose their mnemonic accessibility. Such basic learning procedure underlies the development of both intramodular and intermodular (interface) representations. Importantly, these effects reflect internal statistics of the representations used in processing rather than the frequency with which particular linguistic items appear in a language learner’s environment. Therefore, the statistical learning in MOGUL is defined and constrained by the nature and internal developmental course of the domain-specific “expert systems” rather than statistical patterns in the ambient world.

The statistical mechanism is a primary means of resolving ambiguities at the interfaces. On the initial encounters of non-alternating locatives (e.g., to pour, to spill, to fill) both alternative construals of the locative events are available. What this means in processing terms is that upon hearing the locative constructions (e.g., The waiter poured/spilled/filled…), both Container and Content arguments become sufficiently activated and compete for inclusion into the events’ CS as the plausible recipients of the Agent’s action. Concurrently, the competition proceeds at the SS/CS interface: The dedicated interface processor attempts to find the best match between the verbal direct object (SS) and the highly activated conceptual competitors. The winning structures (i.e., the ones included into the final representational chains), as a result, gain a slightly higher resting activation level which will give them a head start in future events. The frequency with which a given mental structure wins competition and is included in the PS↔SS↔CS chains determines how strongly a particular locative form (PS↔SS) is associated

---

4 The term internal statistics reflects a subjective character of a language learner’s linguistic experiences. It signifies the relationship between frequency of representing/constructing linguistic units (e.g., morpheme, word, phrase, clauses, etc.) during processing and changes in their resting level of activation (i.e., availability of these units for future processing events). On the other hand, external statistics of linguistic units is a convenient abstraction derived from a collection of linguistic utterances produced by speakers of a language.
with a specific semantic construal (CS). This, in turn, determines how readily this form can be reinterpreted/used in the unintended sentence frame. The more strongly a verbal form becomes associated with a given construal, the less compatible it is with the alternative meaning. For example, the highly frequent locatives such as to pour and to dump become strongly associated with the caused motion interpretation (pour/dump water into the bowl) and sound unnatural in the alternative realization (pour/dump the bowl with water). On the other hand, for the less frequent verbs, such as to drizzle and to spill, both construals remain relatively available and the unintended construal/syntactic expression (e.g., drizzle pizza with oil) appears to be more tolerable. In sum, the “use it or lose it” principle adopted in the MOGUL framework provides the statistical means for a resolution of the interface ambiguities: The highly accessible representations make the alternative competitors less accessible.

The fine-tuning of the form-meaning correspondences is also affected by the statistical learning processes within the dedicated “expert systems.” The main function of any module is to compose representations from the activated elements. These composites then can be used for the creation of more complex structures, which, in turn, can serve as the building blocks for even more complex representations, and so on. There are no a priori criteria for how coarse or fine-grained the mental structures must be to be used in processing. The deciding factor in what mental composites can be integrated into a sentence representation is the frequency of their successful use in prior processing. The more often a composite item is accessed and successfully used in representational chain, the more likely it is to be selected on the next processing event. Such chunk-based statistical learning facilitates the processes of linguistic communication by prepackaging “useful” building blocks and making them available for processing (for a review of chunk-based statistical learning, see Perruchet & Pacton, 2006). The highly accessible structures
linked by the corresponding interface representations streamline the linguistic processing and in
doing so reduce ambiguities of the non-alternating locatives and thus diminishing the chances of
their overextension.

Summary

The central thesis of the current account is that the overgeneralizations and retreat from
the overly general patterns of locatives can be best understood as a language processing
phenomenon. They are natural byproducts of the mental processes underlying learners’ attempts
to understand and produce the target constructions. Overgeneralizations are instances of
developmental optionality: availability of unintended semantic senses of the locative verbs
during language comprehension or production (Truscott, 2006). Retreat from the erroneous
patterns involves learning precise semantic distinctions expressed by locative verbs as well as the
statistically-driven strengthening of the appropriate associations between the linguistic forms and
their meanings.

The current account assumes that L1 and L2 learners rely on the same learning
mechanisms and processes in the acquisition of locative constructions. In contrast to Bley-
Vroman and Joo (2001), Joo (2003) and Ambridge and Brandt’s (2013) proposals, it holds that
L2 and L1 learners are not categorically different, even though the quality and quantity of L2
learners’ linguistic experiences as well as L1 representations are likely to affect their ultimate
attainment. Also, the proposal views the learning process as a gradual growth of the target
patterns rather than a switch of the parametric settings advocated by Juffs (1996). The statistical
learning underlies all constructional growth but relies on the internal statistics of the
representational use rather than distributional properties of the environmental input and is
reflected in the mnemonic availability of winning representations rather than the “inferences-
from-absence” or suppression of unused mental structures suggested by Ambridge & Brandt (2013).

In sum, the semi-regularity of the locative constructions can be thought of as a balancing act between human’s flexibility in conceptualizing complex, dynamic events and a tendency towards least ambiguous expression of the meanings in language. Correspondingly, the development these constructions can be thought of as the process of striking the right balance.

The current empirical study

Rethinking the problem of locative constructions in processing terms brings applied linguists closer to an understanding of how language learners’ moment-by-moment linguistic experiences might relate to their knowledge of semi-regular patterns. In its current formulation, however, this account remains quite general and describes the locative phenomena in broad strokes reflecting a paucity of empirical evidence.

In the current project I reexamine L2 learners’ perception of the target structures empirically. The aim is to obtain accurate and unbiased experimental evidence necessary for a resolution of the inconsistencies of the previous research as well as an advancement of testable hypotheses within the processing-based framework advocated by the current project. The projects’ overall goal is to identify patterns in L2 learners’ performance and explore how these patterns vary across the L2 developmental span. Experiment 1 investigates L2 learners’ acceptability judgements. In this experiment, I collect empirical evidence directly comparable to the existing reports. Unlike any prior study, however, Experiment 1 incorporates detailed statistical analyses of L2 learners’ performance comparing their ratings to native speakers’ judgments as well as assessing L2 participants’ judgements in both absolute (accept vs. reject) and preferential (content-oriented vs. container-oriented frame) terms. Experiment 2 addresses
L2 learners’ perception of the locative constructions from a processing perspective. It is designed to rectify shortcomings associated with research methodologies relying on metalinguistic judgments, and, complementarily to Experiment 1, is set up to discover patterns in L2 learners’ real-time auditory comprehension of the target structures. Experiment 2 evaluates participants’ capacity to anticipate verbal arguments in the auditory stream using visual world eye-tracking paradigm (for review, see Altmann, 2012; Huettig, Rommers, & Meyer, 2011). This methodology relies on listeners’ tendencies to look at the visual images (under)mentioned in the speech stream, and, thus, allows to infer from the eye-tracking recordings how participants process the linguistic stimuli in real time. In particular, the methodology makes it possible to test whether L2 learners overcome overgeneralizations by examining how likely they are to predict (i.e., look at) the right kinds of arguments upon hearing non-alternating locative verbs (i.e., Container images for container-only and Content images for content-only locative constructions). Jointly, the two experiments serve the goal of obtaining detailed and unbiased empirical evidence of L2 learners’ processing and representation of locative verbs.
CHAPTER 2: EXPERIMENT 1

Experiment 1 investigated L2 learners’ perception of the locative verb constructions. Mandarin Chinese speaking L2 English learners of low, intermediate, and high proficiency levels and a comparison group of native English speakers took part in an acceptability judgment task contributing their ratings of the target constructions. The main purpose of the experiment was to identify patterns in L2 learners’ acceptability judgments, and in such a way obtain the empirical evidence necessary for a resolution of the apparently contradictory proposals about development and attainment of L2 learners’ knowledge of the locative verbs’ argument structure. In order to gain a detailed picture of the experimental findings, the study’s analysis included a comparison of L2 learners’ judgments to the ratings of the L1 group, along with an examination of the participants’ acceptability judgments within each experimental group in both absolute (accept vs. reject) and preferential (content-oriented vs. container-oriented frame) terms. Also, the study examined cross-sectionally potential developmental trends in L2 learners’ ratings of the locative constructions containing non-alternating verbs and assessed how lexical frequencies of the target locative verbs affect participants’ acceptability judgments. Therefore, Experiment 1 provided a comprehensive assessment of L2 learners’ perception of the locative verb constructions needed for an understanding of how L2 learners deal with the task of learning the semi-productive linguistic patterns. The following research questions guided Experiment 1:
1. Do L2 learners of high, intermediate and low proficiency demonstrate target-like performance on the acceptability judgment task?
   a. Do they accept grammatical and reject ungrammatical locative constructions?
   b. Do they match in their rating performance with the English native group?
   c. Do they demonstrate preferences for the grammatical locative constructions?

2. Do L2 participants’ ratings of the non-alternating locative sentences vary with proficiency?

3. How does lexical frequency of the locative verbs affect L2 participants’ ratings of the target non-alternating locative sentences?

Even though Experiment 1 is exploratory in its nature, the study’s findings are expected to clarify validity of the alternative proposals concerning development and attainment of the locative constructions in L2. Specifically, if L2 learners are not capable of acquiring the fine-grained semantic distinctions implicated in argument realization of the non-alternating locatives, as suggested by Bley-Vroman and Joo (2001), L2 learners are expected to treat uniformly the non-alternating verbs as alternating: They are expected to systematically differ from L1 speakers accepting the ungrammatical constructions (e.g., *spilled the t-shirt and *stained the wine) and have no statistically significant preferences for the alternative realizations of the locative verbs (spilled wine vs. *spilled the t-shirt or *stained the wine vs. stained the t-shirt). Neither are they expected to demonstrate any cross-sectional developmental trends in acquisition of the locative non-alternants. Alternatively, if learning locative verbs is the matter of L1 transfer, as proposed by Juffs (1996), a different pattern of results should appear. In accordance with this position, all groups of L2 learners are predicted to match with L1 speakers in their ratings of the content-only locatives: They are expected to accept grammatical and reject ungrammatical content-only
sentences, and, correspondingly, show statistically significant preferences for the grammatical realization of these locatives (e.g., spilled wine vs. *spilled the t-shirt). At the same time, L2 learners are expected to differ from L1 English speakers in their judgments of the container non-alternating constructions and accept both grammatical and ungrammatical container-only sentences (*stained the wine and stained the t-shirt). The developmental trends are primarily expected for this subset of the locatives with low proficiency L2 learners showing no preferences for the two alternative argument structures. Finally, some tentative predictions can be made based on the processing account advocated by the current project. Therefore, if L2 learners’ perception of the target structures can be best understood as a processing phenomenon, lower proficiency L2 speakers are expected to systematically overextend the non-alternating locatives: diverge from L1 participants by accepting ungrammatical constructions (e.g., *spilled the t-shirt and *stained the wine) and display no preferences for the grammatical sentence frames (spilled wine vs. *spilled the t-shirt or *stained the wine vs. stained the t-shirt). Correspondingly, the developmental trends are expected to show up as a gradual decrease of the acceptance rates for the ungrammatical sentences with high proficiency L2 speakers matching L1 speakers’ rating performance.

Given that there is no empirical evidence concerning effects of lexical frequency of the target verbs on the development and attainment of grammatical intuitions in L2, two rather broad alternative hypotheses (besides the null hypothesis) are viable. First, it is possible that the frequency effects may be observed across all proficiency groups of L2 speakers reflecting relative frequency of the verbs in the target language. Alternatively, the effects might appear only at the advanced levels of L2 proficiency when the learners pass a certain threshold of occurrences and can be guided by the statistical information.
Method

Participants

Native speakers of English. Fifty native speakers of English (6 males and 44 females), undergraduate students at a large university in the United States participated in the study. The mean age of the participants was 21.22 years ($SD = 3.54$), ranging from 18 to 35 years. All participants reported taking foreign language classes or learning languages other than English outside formal settings. Fifteen participants reported attaining intermediate (or above) proficiency level in a foreign language (n=8 in Spanish; n=2 in Korean; and n=1 in each of the following languages: Bosnian, Cantonese, French, Japanese and Polish). All participants were volunteers and received extra course credit or $15 monetary compensation for their participation in the study. The participants’ English proficiency was assessed using the LexTALE test (see Materials section for details) (Lemhöfer & Broersma, 2012). The mean proficiency of the L1 English group was 85.99% ($SD = 9.08$), ranging from 67.5% to 100%.

L2 speakers of English. One hundred and two native speakers of Mandarin Chinese (65 females and 37 males), second language learners of English ranging in age from 18 to 49 years ($M = 23.02$, $SD = 5.40$) took part in the study. The L2 participants were undergraduate or graduate university students. They were volunteers and received $15 compensation for their participation.

The L2 participants were divided into three proficiency groups based on their performance on the LexTALE task. The lower proficiency group (hereafter, Low L2 learners) consisted of 36 participants (23 females and 13 males) with the LexTALE scores ranging from 37.5% to 56.25% ($M = 50.38$%, $SD = 4.24$%). Twenty-eight Low L2 learners were undergraduate students and eight participants were studying towards their graduate degrees.
intermediate English proficiency group (Intermediate L2) included 35 participants, 9 females and 16 males. Thirty Intermediate participants were undergraduate, and five were graduate university students. Their mean English proficiency score was 62.36 % (SD = 3.30) and ranged from 57.5% to 67.5%. The upper level English proficiency group (High L2 learners) consisted of 31 Chinese native speakers (23 females and 8 males) whose LexTALE scores ranged from 68.75% to 87.5% with a mean value of 77.34% (SD = 5.75). Twenty-two High proficiency L2 learners were undergraduate L2 speakers. Table 2 summarizes demographic information for each proficiency group of L2 learners, including their age, age of English learning onset, age of arrival and length of residence in the U.S., LexTALE scores, and percentage of self-reported daily English use.

Table 2. The demographic information for Low, Intermediate, and High proficiency L2 groups

<table>
<thead>
<tr>
<th>L2 groups</th>
<th>Low</th>
<th>Intermediate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Age (in years)</td>
<td>21.64</td>
<td>4.22</td>
<td>22.40</td>
</tr>
<tr>
<td>English learning onset (in years)</td>
<td>9.75</td>
<td>2.70</td>
<td>9.60</td>
</tr>
<tr>
<td>Arrival to the U.S. (in years)</td>
<td>19.51</td>
<td>3.50</td>
<td>20.66</td>
</tr>
<tr>
<td>Residence in the U.S. (in months)</td>
<td>20.42</td>
<td>14.54</td>
<td>19.51</td>
</tr>
<tr>
<td>LexTALE score (in percent)</td>
<td>50.38</td>
<td>4.24</td>
<td>62.36</td>
</tr>
<tr>
<td>Percentage of English daily use</td>
<td>36.39</td>
<td>16.06</td>
<td>40.29</td>
</tr>
</tbody>
</table>

A one-way analysis of variance, comparing the experimental groups’ LexTALE scores, demonstrated that the four groups of participants (English L1, Low L2, Intermediate L2, and
High L2) differed significantly from each other \((F (3, 148) = 248.1, p < .0001)\), and thus, occupied distinct English proficiency bands.

**Materials**

_**Sentence Stimuli.**_ Forty-eight sentences containing locative verbs were the primary targets for the acceptability judgment task. To create the test items, I first selected 24 verbs following the four-way classification of the locative verbs proposed by Pinker (1989). Six content-only, six container-only non-alternating verbs and two six-item sets of alternating verbs (six content-alternating and six container-alternating) comprised the target locative verbs’ set. The selected items are presented in Table 3.

**Table 3. The target locative verbs used in the study**

<table>
<thead>
<tr>
<th>Argument Realization Patterns</th>
<th>Primary Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Content</td>
</tr>
<tr>
<td>Alternating</td>
<td>spread</td>
</tr>
<tr>
<td></td>
<td>pile</td>
</tr>
<tr>
<td></td>
<td>spray</td>
</tr>
<tr>
<td></td>
<td>splash</td>
</tr>
<tr>
<td></td>
<td>smear</td>
</tr>
<tr>
<td></td>
<td>sprinkle</td>
</tr>
<tr>
<td></td>
<td>drip</td>
</tr>
<tr>
<td></td>
<td>dump</td>
</tr>
<tr>
<td>Non-Alternating</td>
<td>shake</td>
</tr>
<tr>
<td></td>
<td>spill</td>
</tr>
<tr>
<td></td>
<td>pin</td>
</tr>
<tr>
<td></td>
<td>pour</td>
</tr>
</tbody>
</table>

In order for the same stimuli to be used in both experiments, the target sentences had to satisfy several additional requirements. One selection criterion for the non-alternating verbs was their compatibility for alternative descriptions of locative events. For example, the verbs _to spill_
and *to stain* can be thought of as alternative ways of expressing a single eventuality in which a liquid such as coffee travels across a space landing on a surface of a t-shirt. In such cases, we can either speak of spilling coffee on the t-shirt or staining the t-shirt with coffee. All selected non-alternating verbs in Table 3, form verb pairs denoting alternative conceptualizations of the same space-time relationship among locative arguments.

I extracted frequency information for the target verbs from the Corpus of Contemporary American English (Davies, 2008). A one-way ANOVA comparing log transformed lexical frequencies of the four categories of the locative verbs revealed no statistically significant differences between the verbs’ types \((F (3, 20) = 0.29, p = .83)\). A summary of the frequencies, jointly with other psycholinguistically relevant information can be seen in Appendix A.

To create the target sentences, I combined each verb with an animate noun, representing an agent of a locative sentence, and two inanimate nouns denoting content and container arguments of a locative event. The content-only and container-only locative verbs were paired with two sets of the content-container arguments (see Materials section of Experiment 2 for additional details and rationale for such a stimuli design).

Each of the 24 verb-arguments combinations was used in both content-oriented and container-oriented sentence frames resulting in 48 target experimental sentences. Each locative verb was used with the same set of arguments in both sentence frames. This was done to ensure that the observed differences in participants’ ratings stem from their perception of how the locative arguments are structurally realized in the sentences, and do not reflect idiosyncratic properties of experimental items that may arise from the different fits of particular arguments with particular verbs. The participants saw a locative verb and three arguments twice presented in the content-oriented and container-oriented constructions. For example, the sentences *The
maid filled tea into the cup and The maid filled the cup with tea served a minimal pair used in the acceptability judgment experiment.

I created two sets of the experimental sentences counterbalancing the pairing of the non-alternating verbs with the content-container arguments. The content-only and container-only verbs appeared exclusively with either argument set A or argument set B in one condition, with the reversed mapping in the counterbalancing version of the stimuli. The sets of the target sentences each containing 48 test items included 36 acceptable and 12 unacceptable constructions. In order to counterbalance the number of grammatical and ungrammatical items and to conceal the focus of the experimental manipulation, I created 24 pairs of distracter items structurally and semantically resembling the targets. The distracter list contained 36 unacceptable and 12 acceptable sentences. A complete set of experimental sentences including target, distractor, and practice items can see seen in Appendix B.

Acceptability Judgments Questionnaire. I created an online acceptability judgment questionnaire using Qualtrics survey software (Qualtrics, 2015). The questionnaire included task instructions, a practice section consisting of eight sentences (two clearly unacceptable, two completely acceptable and four marginally acceptable items), 48 target items, and 48 distractors. The seven-point Likert scale, enabling rating of the sentences in a graded manner, with “completely unnatural” and “completely natural” as the extreme selection options was used to assess participants’ perception of the test items. In order to make the task more intuitive, the Likert scale was accompanied by smiley faces, as done by Blything, Ambridge, and Lieven (2014).

The experimental software delivered the experimental items, one at a time, in a quasi-random order, so that none of the experimental verbs appeared in two sentence frames on two
consecutive trials. On each trial, the participants read a sentence, made their acceptability judgments by clicking on a radio button, and then pressed the “next button” submitting their ratings for a particular test item (Appendix C). The Qualtrics software recorded timing information for each experimental trial, including time of the first mouse click on a Likert scale radio button, time of the last click, a total number of mouse clicks, and the time when the participants submitted their responses by pressing the next button.

**Background questionnaire.** An online background questionnaire, created in Qualtrics, was used to collect data about the participants’ age, gender, current level of education, and L2 learning experiences (Appendix D). The L2 participants answered an additional set of questionnaire items concerning their age of English learning onset, age of arrival and length of residence in an English-speaking country, and percentage of daily English use (Appendix E).

**Proficiency measurement instrument.** To evaluate participants’ English proficiency, I used an online version of the LexTALE: Lexical Test for Advanced Learners of English (Lemhöfer & Broersma, 2012). The LexTALE is an untimed visual lexical decision task. It consists of 60 trials. On each trial participants see a string of letters. Test-takers’ task is to decide whether the letter string is an existing English word or not. The LexTALE score signifies percentage of correct responses for words and non-words used in the test. Lemhöfer and Broersma (2012) conducted a large-scale study in which they compared results of the LexTALE with the Quick Placement Test (QPT) and the Test of English for International Communication (TOEIC). The authors found substantial and highly significant correlations between the LexTALE, QPT and TOEIC suggesting that the LexTALE is a valid measure of L2 proficiency. In addition, Lemhöfer and Broersma (2012) provided evidence that the LexTALE scores were good predictors of L2 participants’ performance on L1-L2 and L2-L1 translation task, as well L2
English lexical decision tasks. The LexTALE has been used in a number of recent studies investigating various aspects of bilinguals’ and L2 learners’ linguistic performance (e.g., Bultena, Dijkstra, & van Hell, 2014; Kaan, Kirkham, & Wijnen, 2016; Khare, Verma, Kar, Srinivasan, & Brysbaert, 2013; Sadri Mirdamadi & De Jong, 2015).

**Vocabulary Knowledge Scale.** To examine the L2 participants’ familiarity with the target vocabulary, the 24 target verbs and 24 distracters were included in an online-based vocabulary knowledge scale survey (see Appendix F). The participants were asked to report their knowledge of the vocabulary on a 4-point Likert scale, adopted from Wesche and Paribakht (1996). The VKS test is a standard measure of L2 learners’ lexical knowledge, which allows efficiently assess whether and how well research participants know the target vocabulary.

**Procedures**

All study participants completed two experiments. The order of their participation was counterbalanced so that half participants in each group first took part in the acceptability judgment study (Experiment 1) and then completed the eye-tracking session (Experiment 2), with the other half participating in the reversed order. There was a two to four day break between the two experiments. Below, I outline the procedures for the participants who started the study with Experiment 1. These procedures differed for people who participated in the acceptability judgment task after the eye-tracking session: The participants began Experiment 1 with the acceptability judgment task, as they had completed the preceding parts on their first session. The VKS test was always the last segment of the study.

The participants were tested in a computer lab in groups of 2 – 12 people. They were first provided with a general overview of the experimental procedures, and then asked to sign an informed consent and fill out the background questionnaire. Then, the participants proceeded to
the LexTALE test. Both oral and written instructions on how to complete the task were given.

After completion of the proficiency assessment task, the participants were asked to work on the acceptability judgment questionnaire. The participants were presented with a scenario in which they were teaching English to a friend. They were asked to imagine that the experimental sentences were produced by their friend, and that their task was to tell him/her whether the sentences sound as well-formed English sentences\(^5\) using the seven-point Likert scale. Eight practice items were presented before the main set of the experimental items. The entire experiment lasted between 40 – 50 minutes.

**Data processing**

The data processing and all statistical analyses were performed by using R Statistical Software (R Development Core, 2015). The acceptability judgment data comprised participants’ ratings of the four categories of locative verbs (content-, container-alternating and content-, container non-alternating) used in the two sentence frames (content-, container-oriented). I coded the acceptability judgment responses numerically from 1 to 7 with 1 representing the “completely unnatural” survey option. The numerical codes were used for the statistical analyses.

To clean the data, I first eliminated “unreasonably fast” submissions on which the participants (most likely) “clicked through” the test items without proper attention to the

---

\(^5\) In an acceptability judgment task, it is unknow whether participants’ ratings reflect specific linguistic features targeted in a study. In this project, participants rated minimal pairs of locative constructions. Each verb was accompanied by the same set of verbal arguments and appeared in two alternative structures. For example, a verb *to fill* was presented in *The waiter filled water into the glass* and *The waiter filled the glass with water* sentence frames. The use minimal pairs ensured that the acceptability ratings signify participants’ perception of the structural realizations of the locative arguments rather than other spurious properties potentially associated with verbal arguments.
experimental task. Given the untimed nature of the experiment, there was a considerable variation in the response timings. Because of this, I could not apply a more principled elimination criterion (such as 2.5 standard deviations below or above a mean) regularly used in reaction time experiments. Therefore, I used a 1.5 second cut-off point as an arbitrary threshold level below which the participants’ submissions were classified as “unreasonably fast”. The submission timing was measured from the appearance of a sentence on a computer screen to the point when a participant pressed the “next button” and on average ranged between 4.54 to 10.21 seconds. The summary of the timing information can be found in Table 13 in Appendix G.

Then, for each experimental group, for each verb type used in a particular sentence frame, I calculated means and standard deviations of the participants’ judgments. The data points below or above 2.5 standard deviations from the respective means were considered as outliers and consequently removed from the data set. Finally, I used L2 learners’ VKS responses in which they indicated either (a) never seeing the target verb before, or (b) seeing the verb before but not knowing its meaning, to discard the acceptability judgments for the sentences containing the unfamiliar words.

An additional data loss occurred due to a recording failure for one participant in the High proficiency L2 group. The final dataset included 74.65% of the originally collected data for Low L2 group, 74.17% for Intermediate L2, 78.02 % for High L2, and 97.88% for L1 experimental group. The detailed summary of the data reduction can be found in Table 14 (Appendix H).

**Statistical Analyses**

**Correlational Analysis.** The initial examination of the experimental outcomes involved a comparison of the acceptability judgments of L1 and L2 groups using a correlational analysis outlined in Ambridge and Brandt (2013). This served as a primarily descriptive device
summarizing findings of the study. I computed difference scores by subtracting the ratings of the test items in the container-oriented frame (e.g., *The bartender poured the glass with water) from the ratings of their content-oriented counterparts (The bartender poured water into the glass). The difference scores signify participants’ preferences for content-oriented constructions and are generally expected to be in a positive range for content-only verbs (e.g., to pour, to spill), take negative values for container-only locatives (e.g., fill, to stain), and cluster around zero for alternating verbs (e.g., to spread, to load). The uniformly computed difference scores (hereafter, “uniform differences”), calculated for each locative verb, allow collapsing acceptability ratings across all verb types and, thus, enable overall comparisons of L1 and L2 groups. The correlational analysis assessed the relationship between the mean “uniform differences” demonstrated by L1 group and those of each group of L2 learners.

Logistic regression analysis. To address the first research question, I performed a logistic regression analysis investigating several interrelated but separable aspects of the participants’ performance on the experimental task. First, I used the analysis to examine how likely the study participants were to judge as acceptable the experimental sentences. Second, I compared the rating behavior of the native speakers and L2 English learners, to determine whether the L2 groups’ acceptance rates differed from those of L1 participants. Finally, I investigated whether the participants were more likely to accept the locative verbs in one sentence frame over the other, looking for the within-group preferences across different verb types. All these aspects of the participants’ judgments were examined within a single statistical model that I ran separately for each category of the locative verbs.

In order to perform the logistic regression analysis, I transformed participants’ seven-point scale responses into a binary outcome variable (i.e., “reject” vs. “accept”) by collapsing
ratings “1”, “2”, and “3” into the “reject” category, and, correspondingly, defining the “5”, “6” and “7” scores as belonging to the category “accept”. The trials on which the participants selected “4” were removed from the dataset (see Appendix I for a summary of data reduction).

The models estimated likelihood of accepting the test sentences using English Proficiency (L1 speakers, High L2, Intermediate L2 and Low L2), and Sentence Frame (Content-Oriented and Container-Oriented) as well as the interaction between them as the predictor variables. The L1 group served as a primary reference category for the English Proficiency independent variable. The ungrammatical constructions (for non-alternating verbs) or the “secondary semantics” sentences (for alternating) were used as the baseline categories for the Sentence Frame predictor. Thus, for example, the intercept for the model predicting acceptance of the content-only verb in the content-oriented (e.g., *pour water into the glass*) vs. container-oriented (e.g., *pour the glass with water*) constructions was the likelihood of accepting the ungrammatical structures by the native speakers’ group. Consequently, all other coefficients of the model were interpreted in relation to the chosen baseline: the effect of the Sentence Frame as the likelihood of accepting grammatical structures (*pour water into the glass*) by the L1 group, and the effects of the English proficiency (High L2, Intermediate L2, and Low L2) as the likelihood of accepting the ungrammatical constructions by the corresponding L2 groups in relation to the likelihood of accepting these items by the native speakers. To compute the additional estimates of interest (i.e. intercept and Sentence Frame values for each proficiency group of L2 learners), I releveled the models’ reference category for the English Proficiency predictor variable.

The initial model fitting revealed a problem known as perfect prediction or complete separation arising in such cases when a predictor variable is perfectly associated with a single
level of the outcome variable. This issue leads to a failure of a computational algorithm to accurately estimate models’ parameters (for further detail, see Gelman et al., 2015). To solve this problem, I used a Bayesian Generalized Linear Modeling procedure (i.e., bayesglm function from “arm” package) which circumvents the separation problem by correcting model’s biases (Gelman, Jakulin, Pittau, & Su, 2008).

**Mixed-effect regression analysis.** To answer the second and third research questions, I used mixed-effect modeling (Baayen, Davidson, & Bates, 2008). Mixed-effect models have a number of advantages over the traditional repeated measures ANOVAs enabling generalization of statistical results for both participants and items within a single analysis. The outcome variable was a preference score, calculated for each pair of test items (for each participant) by subtracting a rating of an ungrammatical (dispreferred) item from a rating of its grammatical (preferred) counterpart. Thus, for instance if a participant rated the sentence *The maid filled the glass with water* “6”, and the sentence *The maid filled water into the glass* “2”, the preference score was “4”, which represented the degree of a participant’s preference for the grammatical item.

Given that the primary focus of the analysis was the variance of L2 participants’ “preference scores” across different proficiency levels, as well as the role of lexical frequencies in the observed preferences, the analysis concentrated on the non-alternating locatives for which the variance and frequency effects are expected on theoretical grounds.

I ran a mixed-effect model which included Verb Type (content-only vs. container-only), Proficiency Level (Low L2, Intermediate L2, and High L2), and Lexical Frequency, as well as all interactions between these predictor variables. All independent categorical variables were dummy coded, therefore, the model estimates represented simple main effects. The random
effect structure included by-subject and by-item random intercepts together with the by-subject random slopes for the Verb Type predictor. To obtain significance values for the predictors and their interactions, I used the Kenward-Roger approximation procedure implemented in the `pbkrtest` package (Ulrich Halekoh, 2014).

Results

Overview of the experimental findings

Before delving into details, I offer a “bird eye view” of the experimental findings using a correlational analysis as a primarily descriptive device summarizing the overall patterns found in the study. Figure 2, Figure 3, and Figure 4 plot mean “uniform differences” for the L1 English group against the “uniform difference” scores for Low L2, Intermediate L2, and High L2 experimental groups, respectively.

The green diagonal line in each graph represents an identity line or line of equality. If the compared groups’ “uniform differences” were identical, they would fall exactly on the line of equality, signifying perfect match between the scores of the two groups. The line of best fit (red dotted line) represents the overall trend in the relationship of the “uniform difference” scores. Figures 1, 2, and 3 depict several important patterns in the data. First, as evident from the graphs, the ratings of the L1 and L2 groups are neither perfectly aligned nor are they totally unrelated.
Figure 2: “Uniform differences” (preference for content- over container-oriented uses of the locative verbs by categories) for L1 (x axis) and Low L2 (y axis) speakers of English.
Figure 3. “Uniform differences” (preference for content- over container-oriented uses of the locative verbs by categories) for L1 (x axis) and Intermediate L2 (y axis) speakers of English.
Figure 4. “Uniform differences” (preference for content- over container-oriented uses of the locative verbs by categories) for L1 (x axis) and High L2 (y axis) speakers of English.

Despite strong positive correlations between native speakers’ and second language learners’ ratings (L1 and Low L2, $r = .73, p < .0001$; L1 and Intermediate L2, $r = .85, p < .0001$; and L1 and High L2, $r = .86, p < .0001$), the L2 learners generally overextended both content and container non-alternating locative verbs demonstrating smaller preferences for the grammatical constructions. For instance, a relatively large difference score for the content-only verb to shake demonstrated by the L1 group ($M = 4.37, SD = 1.67$) was systematically lower for all groups of
L2 speakers (Low L2, $M = 0.06, SD = 2.39$; Intermediate L2, $M = 1.38, SD = 2.47$; and High L2, $M = 2.71, SD = 2.46$). The second important pattern manifested in Figures 1, 2, and 3 is that the line of best fit grows steeper for the higher proficiency L2 groups, gradually approaching the identity line. The observed trend, jointly with the correlation coefficients reported above, suggest that the “uniform differences” of L2 learners do not vary randomly, but tend to become more native-like with higher proficiency in English. The third notable pattern observed across all graphs is that the L2 learners’ “uniform differences” for the alternating verbs tend to take positive values. Thus, for a majority of the content-alternating (blue squares) and container-alternating (purple triangles) verbs, L2 participants preferred their content-oriented realizations. Finally, within each category of the locative verbs there is a considerable scatter of the “uniform difference” values, suggesting that it is not the case that all verbs belonging to a particular category of the locative verbs evoke identical preferences. For example, among six content non-alternating verbs (red circles) rated by the High proficiency L2 learners (Figure 3), the verb *to pour* and *to drip* received substantially different preference scores. Whereas the High L2 speakers strongly preferred the content-oriented use of the verb *to pour* ($M = 3.23, SD = 2.45$), with the “difference score” practically matching the rating of the L1 group ($M = 3.46, SD = 2.04$), they demonstrated only marginal preference for the verb *to drip* ($M = 0.75, SD = 1.88$), contrasting from the native speakers’ judgment ($M = 2.94, SD = 2.14$).

In summary, the initial data exploration suggests a close correspondence between native speakers’ and L2 learners’ ratings of the locative verbs. At the same time, the observed patterns of divergence indicate greater acceptability of the ungrammatical non-alternating constructions, as well as general preferences for the content-oriented realization of the alternating verbs by the L2 experimental groups. In addition, the analysis demonstrates an apparent “developmental
“trend” across L2 learners’ groups, which, however, varies for the specific items within a particular category of the locative verbs.

**Acceptability of the locative constructions**

To answer the question concerning the target-like performance of the L2 speakers on the acceptability judgment task, I consider several interrelated but separable aspects of the participants’ rating performance. First, I present native speakers’ ratings, in this way establishing the benchmarks against which the L2 participants’ judgments can be assessed. Then, I explore the L2 learners’ judgments and compare them to the native speakers’ outcomes. Finally, for each experimental group, I compare the alternative argument realizations of the locative verbs, looking into participants’ preferences across different categories of the locatives.

All reported statistics are based on the results of the logistic regression analysis performed separately for each category of the locatives (see Appendix J for a complete summary of the models). For the clarity of exposition, I introduce the regression estimates following the presentation logic outlined above.

*Interpreting logistic regression coefficients.* The interpretation of logistic regression coefficients differs from interpretation of general linear models (for further details, see Agresti, 2013; Gelman & Hill, 2007). The model’s parameters appear in a *logit* metric and are not readily interpretable (besides general directions of the statistical results). A more intuitive alternative is to reintroduce the statistical results in terms of odds ratios (OR) by exponentiating the logit coefficients. Odds ratios are relative measures of experimental effects. They represent effects of particular predictor variables (on the observed outcomes) in relation to an effect of a base line category and are interpreted on a multiplicative scale. For example, by exponentiating a hypothetical logit estimate 1.3 for a “variable,” we receive 3.67 on the odds ratio scale, and can
interpret it as indicating that the likelihood of an event of interest to occur is 3.67 times higher (given the “variable1”) in comparison to the likelihood of the event’s occurrence for the baseline category (“variable0”).

A critical feature of the current logistic analysis is that the intercepts of the models are fully interpretable. Given that the logistic models predict likelihood of accepting the locative sentences, the intercepts provide information about likelihood of accepting the reference category items: ungrammatical locatives of the non-alternating verbs, and, for the alternating verbs, locative constructions in the “secondary semantics” sentence frames. Correspondingly, the statistical significance tests for the intercepts evaluate the baseline category effects relative to a chance level (i.e. “0” on the logit scale or “1” in the odds ratio terms), providing the statistical means for assessing participants’ acceptability of the baseline category items.

**Examining target-like performance.** Figures 5, 6, 7, and 8 graphically summarize the descriptive statistics for participants’ performance on the acceptability judgment task, representing average acceptance rate of the target locative sentences in the content- and container-oriented constructions by the experimental groups.

Examining the intercept values of the logistic models, we can see that the L1 group accepted both reference category items of the alternating locative verbs: The content-alternating locative verbs in the container-oriented sentence frame \( (M = 93.40, SD = 24.87, \beta = 2.66, SE = 0.23, \ p < .0001, OR = 14.29) \), as well as the container-alternating locatives in the content-oriented construction \( (M = 100.00, SD = 0.00, \beta = 14.29, SE = 0.99, \ p < .0001, OR = 706.55) \) were strongly accepted by the native speaking participants.
Figure 5. Mean acceptance rate of the container-alternating locative verbs by the experimental groups. Error bars represent standard errors.

Figure 6. Mean acceptance rate of content-alternating locative verbs by the experimental groups. Error bars represent standard errors.
Figure 7. Mean acceptance rate of the container non-alternating locative verbs by the experimental groups. Error bars represent standard errors.

Figure 8. Mean acceptance rate of the content non-alternating (right) locative verbs by the experimental groups. Error bars represent standard errors.
The content non-alternating sentences in the container-oriented frame (e.g., *The bartender poured the glass with water*) were rarely accepted by the L1 participants \((M = 26.62, \ SD = 44.28, \ \beta = -0.97, \ SE = 0.13, \ p < .0001, \ OR = 0.38)\) indicating their dispreference of the ill-formed locative structures. At the same time, the ungrammatical container-only constructions (e.g., *The maid filled water into the glass*) were as likely to be accepted as rejected with the acceptance rate at the chance level \((M = 47.84, \ SD = 50.04, \ \beta = -0.006, \ SE = 0.12, \ p = .62, \ OR = 0.94)\). Similarly to L1 participants, all groups of L2 speakers accepted the content-alternating verbs in the container-oriented constructions \((\text{High L2}, \ M = 77.40, \ SD = 41.97, \ \beta = 2.42, \ SE = 0.29, \ p < .0001, \ OR = 11.25; \ \text{Intermediate L2}, \ M = 66.13, \ SD = 47.52, \ \beta = 2.21, \ SE = 0.26, \ p < .0001, \ OR = 9.12; \ \text{Low L2}, \ M = 70.73, \ SD = 45.69, \ \beta = 1.87, \ SE = 0.23, \ p < .0001, \ OR = 6.49)\), as well as the container-alternating locatives in the content-oriented frames \((\text{High L2}, \ M = 89.43, \ SD = 30.87, \ \beta = 0.52, \ SE = 0.18, \ p = .004, \ OR = 1.68; \ \text{Intermediate L2}, \ M = 90.26, \ SD = 29.75, \ \beta = 0.68, \ SE = 0.18, \ p = .0003, \ OR = 1.97; \ \text{Low L2}, \ M = 86.54, \ SD = 34.24, \ \beta = 0.89, \ SE = 0.19, \ p < .0001, \ OR = 2.43)\). Therefore, in parallel to the native speaking participants, L2 English learners of high, intermediate, and low proficiency levels were significantly more likely to accept (in comparison to the chance level) such locative sentences as *John spread the toast with butter* and *Susan loaded clothes into the suitcase*, demonstrating the expected acceptability rating patterns.

Unlike locative alternants, the non-alternating constructions were perceived quite differently by the L2 participants. In contrast to L1 speakers, High L2 and Intermediate L2 groups were equally likely to accept or reject content-only sentences such as *The bartender poured the glass with water* \((\text{High L2}, \ M = 45.38, \ SD = 49.98, \ \beta = -0.16, \ SE = 0.17, \ p = .35, \ OR = 0.85; \ \text{Intermediate L2}, \ M = 53.24, \ SD = 50.08, \ \beta = 0.12, \ SE = 0.17, \ p = .47, \ OR = 1.13)\), and
Low L2 participants were statistically more likely to accept such constructions ($M = 61.39$, $SD = 48.84$, $\beta = 0.43$, $SE = 0.16$, $p = .007$, $OR = 1.54$). Finally, all groups of L2 speakers uniformly accepted the ungrammatical container non-alternants such as *The maid filled water into the glass* (L2 High, $M = 62.09$, $SD = 48.68$, $\beta = 0.50$, $SE = 0.16$, $p = .002$, $OR = 1.65$; Intermediate L2, $M = 62.66$, $SD = 48.52$, $\beta = 0.51$, $SE = 0.16$, $p = .002$, $OR = 1.67$; Low L2, $M = 64.78$, $SD = 47.92$, $\beta = 0.60$, $SE = 0.16$, $p = .0002$, $OR = 1.82$).

In sum, the L2 learners of English did not demonstrate the (theoretically assumed) target-like performance on the acceptability judgment task. Although, similarly to native speakers, all L2 experimental groups accepted alternating constructions, they failed to reject the ungrammatical locative structures. Importantly, the participants’ divergences from the expected norms formed consistent patterns across the proficiency levels and differed for the content and container non-alternating verbs. The ungrammatical content non-alternating constructions, rejected by the native speakers, received undetermined evaluations from both High and Intermediate L2 learners, and were accepted by the Low proficiency L2 speakers. In turn, the ungrammatical container-only sentences, for which L1 speakers showed no definite acceptability judgments, were uniformly accepted by all L2 speaking groups.

**Comparing L1 and L2 acceptability judgments.** All L2 speaking groups consistently differed from the native speakers in their rating performance. In comparison to the L1 participants, they demonstrated significantly lower acceptance rates for the content-alternating reference structures (High L2, $\beta = -2.16$, $SE = 0.29$, $p < .0001$, $OR = 0.11$; Intermediate L2, $\beta = -1.99$, $SE = 0.30$, $p < .0001$, $OR = 0.14$; Low L2, $\beta = -1.78$, $SE = 0.30$, $p < .0001$, $OR = 0.17$), as well as for the container-alternating locatives in the content-oriented sentence frame (High L2, $\beta = -4.09$, $SE = 1.02$, $p < .0001$, $OR = 0.02$; Intermediate L2, $\beta = -4.33$, $SE = 1.01$, $p < .0001$, $OR = 0.01$).
$\text{OR} = 0.01; \text{Low L2}, \beta = -4.69, SE = 1.01, \ p < .0001, \text{OR} = 0.009).$ Thus, whereas the L1 group demonstrated total (for the container alternants) or nearly total (for the content-alternants) acceptance of the grammatical locative structures, the L2 groups showed at least some degree of uncertainty in their judgments.

The L2 learners showed the opposite pattern of divergences in their judgments of the ungrammatical non-alternating locatives. In comparison to the L1 group, all L2 participants were significantly more likely to accept the ungrammatical content non-alternating (High L2, $\beta = 0.77$, $SE = 0.22, \ p < .0001, \text{OR} = 2.16$; Intermediate L2, $\beta = 1.09, SE = 0.21, \ p < .0001, \text{OR} = 2.97$; Low L2, $\beta = 1.42, SE = 0.21, \ p < .0001, \text{OR} = 4.14$), as well as container non-alternating sentences (High L2 $\beta = 0.54, SE = 0.20, \ p = .008, \text{OR} = 1.72$; Intermediate L2, $\beta = 0.56, SE = 0.20, \ p = .0005, \text{OR} = 1.75$; Low L2, $\beta = 0.65, SE = 0.20, \ p = .0001, \text{OR} = 1.92$). Therefore, all L2 speaking groups consistently demonstrated greater tolerance of the grammatical violations.

**Examining preferences in the acceptability judgments.** All groups of participants showed statistically significant preferences for the well-formed sentences containing non-alternating locative verbs. The within-group comparisons of the content-only locatives in the content-oriented frame (e.g., *The bartender poured water into the glass*) with the container-oriented realizations of these verbs (e.g., *The bartender poured the glass with water*) demonstrated uniform preferences for the former constructions across all experimental groups of participants (L1, $\beta = 6.41, SE = 0.80, \ p < .0001, \text{OR} = 607.89$; High L2, $\beta = 3.43, SE = 0.47, \ p < .0001, \text{OR} = 30.88$; Intermediate L2, $\beta = 1.89, SE = 0.30, \ p < .0001, \text{OR} = 6.62$; Low L2, $\beta = 1.27, SE = 0.27, \ p < .0001, \text{OR} = 3.56$). Similarly, the participants across all proficiency levels preferred grammatical realizations of the container non-alternating verbs (e.g., *The maid filled the cup with tea*) over their ungrammatical counterparts (e.g., *The maid filled tea into the cup*),
(L1, \( \beta = 5.42, SE = 0.75, \ p < .0001, \ OR = 225.88 \); High L2, \( \beta = 2.47, SE = 0.39, \ p < .0001, \ OR = 11.82 \); Intermediate L2, \( \beta = 1.87, SE = 0.31, \ p < .0001, \ OR = 4.49 \); Low L2, \( \beta = 2.21, SE = 0.36, \ p < .0001, \ OR = 9.12 \)). Therefore, even though the English native speakers were as likely to accept as to reject the ungrammatical container-only sentences, and all groups of L2 speakers were more likely to rate these structures as acceptable, all groups participating in the study demonstrated the theoretically expected preferences for the well-formed container non-alternating constructions.

As has been noted in the overview of the experimental findings, the study participants showed a general bias towards content-oriented realization of the alternating locative verbs. The statistical results indicated that the content-alternating verbs in the content-oriented frame such as *John spread butter onto the toast* received significantly higher acceptability ratings in comparison with their container-oriented realizations, such as *John spread the toast with butter* (L1, \( \beta = 2.75, SE = 0.70, \ p < .0001, \ OR = 15.64 \); High L2, \( \beta = 1.59, SE = 0.33, \ p < .0001, \ OR = 4.90 \); Intermediate L2, \( \beta = 1.13, SE = 0.31, \ p = .0002, \ OR = 3.09 \); Low L2, \( \beta = 1.23, SE = 0.33, \ p = .0002, \ OR = 3.42 \)). These systematic preferences were observed across all groups of participants even though all of them accepted both argument realization patterns. Similarly, the content-oriented preferences were observed for the container-alternating verbs. Such sentences as *Susan loaded clothes into the suitcase* were ranked consistently higher than their container-oriented alternatives, such as *Susan loaded the suitcase with clothes*, by all L2 groups (High L2, \( \beta = -1.17, SE = 0.34, \ p = .0006, \ OR = 0.31 \); Intermediate L2, \( \beta = -1.13, SE = 0.31, \ p = .0002, \ OR = 0.32 \); Low L2, \( \beta = -1.04, SE = 0.28, \ p = .0002, \ OR = 0.35 \).
Changes of acceptability preferences across L2 proficiency levels

The second research question inquired whether the L2 participants’ ratings of the locative sentences varied across proficiency levels, looking into the “developmental trends” primarily expected for the non-alternating verbs. Figure 9 shows mean preference scores for the content- and container-only locatives for three groups of L2 participants. As evident from Figure 9, the preferences scores (i.e. preferences for the well-formed argument realizations) for both categories of non-alternating verbs increased with proficiency: Higher proficiency L2 learners demonstrated greater preferences for the grammatical structures. The mixed-effect model analysis revealed statistically significant differences between Low L2 ($M = 1.08, SD = 3.45$) and High L2 ($M = 2.30, SD = 3.13$) experimental groups’ preferences scores for the content-only locatives ($\beta = 1.12, SE = 0.39, \ p = .005$), but not for the container non-alternants (Low L2, $M = 1.43, SD = 2.99$ vs. High L2, $M = 1.95, SD = 2.65$; $\beta = 0.53, SE = 0.37, \ p = .16$).

![Figure 9](image-url)

**Figure 9.** Mean preference scores for content and container non-alternating locative verbs for low, intermediate and high proficiency L2 speakers. Error bars represent standard errors.
These findings suggest that the acceptability ratings of the non-alternating locative verbs do vary across L2 participants of different English proficiency levels: More advanced L2 learners tend to make greater distinctions between grammatical and ungrammatical locatives. The observed developmental trend, however, appears to be more pronounced for the content-only constructions.

**Frequency effects in acceptability judgment of the locative verbs**

The third research question concerned effects of lexical frequency of the target verbs on acceptability of the locative constructions. It was addressed, jointly with the previous research question, by the mixed-effect regression analysis (see Appendix K for a summary of the model). The results of mixed-effect analysis showed that the statistically significant frequency effects were observed exclusively for the High L2 group and differed between two categories of the locative non-alternants \( \text{Proficiency}_{\text{HIGH}} \times \text{Verb Type}_{\text{CONTAINER}} \times \text{Frequency}_{\text{(COCA)}} \), \( \beta = -2.12, SE = 0.73, p = .005 \). The follow up analysis conducted separately for the content-only and container-only locative constructions rated by the High L2 learners revealed that the frequency effects were associated with the preference scores of the content-only verbs \( \beta = 1.63, SE = 0.79, p = .037 \), but did not affect ratings of the container non-alternating verbs \( \beta = 0.11, SE = 0.27, p = .65 \). The statistically significant positive relation between frequency of the content non-alternating verbs and the High L2 speakers’ preference scores suggests that the L2 participants’ distinctions between grammatical and ungrammatical constructions increase with the lexical frequency: the higher frequency, the greater preferences for the grammatical content non-alternating sentences.

Given that the observed frequency effect can be a production of either higher rating of the grammatical structures or lower acceptability of their ungrammatical counterparts (or a
combination of both), the follow up analysis examined the frequency effects using the raw rating scores. Figure 10 displays the relation between lexical frequency of the content non-alternating verbs (x-axis) and the High L2 participants’ acceptability judgment scores (y-axis) for the grammatical (blue) and ungrammatical (red) content-only constructions. The mixed-effect regression analysis modeling the participants’ judgment scores as the function of the Sentence Frame (Content vs. Container) and Lexical Frequency independent predictors (and their interaction) showed that the frequency effects were different for the content-oriented and container-oriented realizations of the content non-alternating verbs (Sentence Frame*Frequency, \(\beta = 1.46, SE = 0.52, p = .009\)): Whereas the ratings of the grammatical sentences (e.g., *The bartender poured water into the glass*) did not vary with the frequency (\(\beta = 0.24, SE = 0.51, p = .66\)), the acceptability scores for the ungrammatical constructions (e.g., *The bartender poured the glass with water*) tended to be lower for higher frequency locative verbs (\(\beta = -1.27, SE = 0.63, p = .04\))
Figure 10. Effect of lexical frequency on High L2 participants’ ratings of content non-alternating locative verbs in content-oriented (blue dots) and container-oriented (red dots) sentence frames

Summary of the experimental results

The participants’ acceptability ratings form consistent patterns across different proficiency levels and categories of the locative verbs. The L2 learners and the comparison group of L1 English speakers demonstrated consistent acceptance of the test items containing alternating locatives. Compared to native speakers, all groups of L2 learners were less likely to accept the alternating constructions demonstrating at least some degree of uncertainty in their acceptability judgments. Even though all L2 participants accepted the alternating locative constructions, each L2 group showed statistically significant preferences for the content-oriented items (e.g., John spread butter onto the toast and Susan loaded clothes into the suitcase) over the container-oriented constructions (John spread the toast with butter and Susan loaded the suitcase with clothes).

The experimental results clearly demonstrate that all groups of participants distinguished between alternating and non-alternating verbs. The L2 learners of low, intermediate and high
proficiency were more likely to accept the ungrammatical sentences when compared to L1
speakers, but, at the same time, showed consistent preferences for the grammatical realizations of
content and container non-alternating locative verbs relatively to the ratings of the
ungrammatical sentences. Table 4 summarizes findings for this subgroup of the locatives. The
examination of the experimental outcomes in the absolute terms (i.e., in relation to a chance
level) reveals consistent patterns in the acceptability judgments. All participants showed high
acceptance rates for the grammatical constructions (means and standard deviations are shown in
square brackets). On the other hand, the ratings of the ungrammatical sentences varied across
proficiency groups forming a clear implicational scale. The content-only ungrammatical
constructions, rarely accepted by the L1 speakers, received uncertain ratings from high and
intermediate L2 learners, and were more likely to be accepted by the low L2 speakers.
Correspondingly, the container non-alternating ungrammatical structures, rated uncertainly by
the L1 groups, were more likely to be accepted by all groups of L2 learners.

Table 4. Percentage of accepting non-alternating constructions by the experimental groups

<table>
<thead>
<tr>
<th></th>
<th>Low L2</th>
<th>Intermediate L2</th>
<th>High L2</th>
<th>L1 Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>Container Only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*stained the wine</td>
<td>64.78 (47.92)</td>
<td>62.66 (48.52)</td>
<td>62.09 (48.68)</td>
<td>47.84 (50.04)</td>
</tr>
<tr>
<td>stained the t-shirt</td>
<td>[94.41 (23.04)]</td>
<td>[91.41 (28.11)]</td>
<td>95.33 (21.16)</td>
<td>[100.00]</td>
</tr>
<tr>
<td>Content Only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*spilled the t-shirt</td>
<td>61.39 (48.84)</td>
<td>53.24 (50.08)</td>
<td>45.38 (49.98)</td>
<td>26.62 (44.28)</td>
</tr>
<tr>
<td>spilled the wine</td>
<td>[84.38 (36.42)]</td>
<td>[88.11 (32.48)]</td>
<td>[96.95 (17.27)]</td>
<td>[100.00]</td>
</tr>
</tbody>
</table>

Notes, Colors of the cells signify acceptance of the ungrammatical structures: green = accept, yellow = uncertain, red = reject
The content-oriented bias is also evident when we examine how L2 participants’ ratings of the ungrammatical constructions vary across proficiency groups: The expected developmental trend, that is, statistically significant decrease in acceptability of the ungrammatical test items is characteristic only for content-only sentences. Similarly, the frequency effects are obvious only for content non-alternating constructions rated by high proficiency L2 learners.
CHAPTER 3: EXPERIMENT 2

Experiment 2 investigated the participants’ online processing of the locative constructions using “visual world” eye-tracking methodology. The experimental groups of native English speakers and L2 learners of low, intermediate, and high English proficiency listened to the target sentences presented concurrently with the visual displays showing entities of the locative events. The eye-tracking equipment recorded the participants’ eye fixations during the unfolding linguistic stimuli. The crux of the “visual world” paradigm methodology is in a tight coordination between linguistic processing and oculomotor behavior: Listeners tend to look at the visual images mentioned in the linguistic stream. Moreover, in many cases the eye-movements reveal anticipation of the upcoming information in a sentence. For instance, when hearing the sentence The boy will eat … participants are more likely to look at an edible visual stimulus (e.g., the cake) prior to its appearance in the auditory stream (Altmann & Kamide, 1999). Experiment 2 examined the verb-based anticipatory processing of the target locative sentences. In essence, the study recast Experiment 1 in processing terms evaluating whether the study participants demonstrated target-like performance in the online comprehension task.

By hypothesis, if participants have knowledge of the argument realization patterns of the target verbs, they are expected to predict the upcoming locative arguments: Upon hearing sentences containing content-only verbs, such as pour, spill, or drip, they should direct their anticipatory eye-movements to the Content images (e.g., water, coffee, juice). Correspondingly, when presented with the container non-alternating locative constructions with such verbs as fill, cover, or stain, the listeners are expected to initiate more looks to the images of Container (e.g., glass, plate, or jeans) before they hear these arguments. Such an experimental design offers several important advantages eliminating or minimizing inherent shortcomings of the
acceptability judgment task. In Experiment 1, the participants’ ratings could be influenced by their response strategies. The respondents could rate the test items as (un)acceptable to maintain the acceptance/rejection balance across the experimental task. Also, the acceptability judgments could reflect the participants’ metalinguistic reasoning, thus, at least to some degree distorting the experimental results. Neither of these factors has a significant effect in the online comprehension task.

Most importantly, however, is that the acceptability judgment task creates a situation in which the ungrammatical test items, even if perceived as sounding somewhat unnatural, still can be interpreted as fully meaningful sentences. The participants reading such sentences as The waiter filled water into the glass or The maid poured the glass with water, could still “repair” or coerce the verbal meanings achieving the fully interpretable messages. In turn, such forced interpretability of the target sentences could result in overestimation of the participants’ acceptance of the ungrammatical structures. Experiment 2 fully eliminates the potential influence of the forced interpretability of the test items by focusing on the participants’ anticipatory language processing. The following research question guided Experiment 2:

*Do native English speakers and L2 learners of English of high, intermediate, and low proficiency levels anticipate the target arguments of the non-alternating locative verbs?*

The experimental findings of the current investigation are expected to provide further evidence needed for an understanding of L2 acquisition of locative structures. Concomitantly, the eye-tracking results are expected to contribute to a resolution of the existing conflicting proposals concerning development and attainment of the locative constructions in L2. In parallel to Experiment 1, the study’s findings can be interpreted along the following lines. If L2 learners are not capable of acquiring the nuances of the locative constructions, as suggested by Bley-
Vroman and Joo (2001), none of the L2 experimental groups will exhibit the anticipatory eye-
movements. If, on the other hand, the native language of L2 learners plays a deciding role in the
acquisition of L2 locatives (Juffs, 1996), all L2 groups will anticipate the correct arguments of
the content-only verbs (i.e., water, coffee, or juice when hearing the verbs pour, spill or drip),
but, at the same time, will fail to do so for the container non-alternating constructions in the early
stages of L2 learning. In this case, the anticipatory processing is expected to emerge in the later
stages of L2 learning. Finally, if L2 growth of the locative constructions can be best understood
in processing terms, the anticipatory eye-movements are expected to correlate with the L2
learners’ proficiency: only more advanced L2 learners of English will demonstrate the
anticipatory processing of the target structures.

Method

Participants

The same groups of native and L2 speakers that took part in Experiment 1 participated in
Experiment 2. All participants reported normal or corrected to normal vision and hearing.

Materials

Sentence stimuli. The sentence stimuli from Experiment 1 were adopted for Experiment
2 with several important modifications. Only grammatical target items were used in Experiment
2. Twelve pairs of the alternating locative constructions as well as four six-item sets of the
grammatical non-alternants comprised the target stimulus pool. The alternating locative stimuli
were six content- and six container-locative alternants realized in both content- and container-
oriented sentence frames. The non-alternating constructions were sentences created by pairing
six content-only and six container-only verbs with each of the two sets of the Content-Container
Arguments shown in Table 5. The pairing of the non-alternating verbs with the two argument
sets ensured that across the study the same visual images served as both targets and competitors. For instance, I paired the verbs pour and fill with the water/glass and tea/cup argument sets resulting in such constructions as …poured the water into the glass, pour the tea into the cup, fill the glass with water or fill the cup with tea. All of the four nouns, therefore, served as the anticipated post-verbal arguments: water and tea for the verb pour, and glass and cup for the verb fill. Such stimuli configuration allowed the use of the same visual displays to measure participants’ eye movement behavior associated with processing both categories of non-alternating locative sentences.

Table 5. The non-alternating locative verbs and two sets of content-container arguments

<table>
<thead>
<tr>
<th>Non-alternating locative verbs</th>
<th>Content/Container Arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Content Only</td>
</tr>
<tr>
<td>drip</td>
<td>cover</td>
</tr>
<tr>
<td>dump</td>
<td>coat</td>
</tr>
<tr>
<td>shake</td>
<td>season</td>
</tr>
<tr>
<td>spill</td>
<td>stain</td>
</tr>
<tr>
<td>pin</td>
<td>decorate</td>
</tr>
<tr>
<td>pour</td>
<td>fill</td>
</tr>
</tbody>
</table>

I created four lists of the sentence stimuli counterbalancing (a) appearance of the alternating verbs in one of the two alternative frames, and (b) pairing of the non-alternating verbs with the alternative argument sets. Thus, half of the alternating verbs appeared in the content-locative frames (e.g. …spread the butter onto the toast, …crammed the clothes into the suitcase), and the other half in the container-locative constructions (piled the table with the dishes, loaded the cart with the vegetables), with the reversed framing in the counterbalancing
condition (e.g., *spread the toast with butter, cramped the suitcase with the clothes* vs. *piled the dishes onto the table and loaded the vegetables into the cart*). Similarly, the content-only and container-only verbs appeared exclusively with either argument set A or argument set B in one condition, with the reversed mapping in the counterbalancing version of the stimuli. In each of the four versions of the stimuli, each verb appeared only once in either content-locative or container-locative constructions (for the alternating verbs) or paired with the argument set A or set B (for the non-alternating verbs). Each participant was presented with only one version of the stimuli. The four versions of the stimuli were equally distributed across the participants in each experimental group, so that overall all verbs equally appeared in each condition.

Twenty-four distractor sentences from Experiment 1 were used in the eye-tracking study. The ungrammatical distractors were revised so that all of them were well-formed English sentences. An additional set of 24 distractors was created in order to maintain 2:1 distractor to target proportion. Three agents, *the cook, the waiter*, and *the manager* were used as the subjects of the experimental sentences and were equally distributed across the targets and distractors in each stimuli set. The complete set of experimental sentence stimuli included 24 alternating, 24 non-alternating, and 48 distractor items (Appendix L).

**Auditory and visual stimuli.** The auditory stimuli were recorded using the NeoSpeech™ TTS On-Demand online service (Neospeech, 2014), and then edited in Audacity audio editing software. All sentence stimuli were synthesized using Julie (female) voice speaking with standard American English accent. Table 6 shows descriptive statistics for the duration of the verb, determiner, and direct object noun segments. One-way analyses of variance comparing the segments’ durations showed that neither verb \( F (3, 20) = 1.04, p = .40 \) nor article \( F (3, 44) = \)
0.29, \( p = .83 \) segments differed statistically across the four categories of the locative verbs.

Marginally statistical differences were found for the noun segment \( (F (3, 44) = 2.48, \ p = .08) \).

Table 6. *The mean duration and standard deviation (in parenthesis) in milliseconds for verb, article, and noun segment of the target auditory stimuli by verb type*

<table>
<thead>
<tr>
<th>Sentence segment</th>
<th>Verb</th>
<th>Article</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content only</td>
<td>670.33 (119.92)</td>
<td>167.08 (46.46)</td>
<td>557.42 (85.98)</td>
</tr>
<tr>
<td>Container only</td>
<td>738.33 (91.96)</td>
<td>170.50 (33.70)</td>
<td>573.75 (82.03)</td>
</tr>
<tr>
<td>Content Alternating</td>
<td>682.66 (135.69)</td>
<td>174.08 (40.88)</td>
<td>482.75 (90.85)</td>
</tr>
<tr>
<td>Container Alternating</td>
<td>627.17 (53.41)</td>
<td>158.75 (46.78)</td>
<td>573.50 (120.72)</td>
</tr>
</tbody>
</table>

A set of color photographs depicting the subjects, and direct and indirect objects of the experimental sentences were used as the target visual objects. All of the photographs were images acquired from Google Images service (see Appendix M for a sample of visual stimuli).

**Experimental presentations.** The auditory sentence stimuli and the corresponding images were used to create the experimental presentations using E-Prime 2.0 with E-Prime Extension for Tobii software (Psychology Software Tools, Pittsburgh, PA). Each visual display included three images. The picture of the subject (agent) always appeared in the middle at the bottom of the screen. The images of the direct and indirect objects (content, containers) were displayed to the left and right, located equidistantly from the center of the screen. The areas of the three images (each 10.20 × 10.15 cm large) served as the visual Regions of Interests (ROIs) investigated in the study. The positions of the content and container objects were randomized across the
experimental trials. The experimental presentations consisted of written and auditory instructions, a calibration session, eight practice and 72 experimental trials. The order of presentation of the experimental slides was randomized across the experiment.

**Procedure**

The eye-tracking study was conducted in a laboratory. Each participant was tested individually. The visual displays were presented on a 23-inch viewing monitor at a resolution of 1920 x 1080 pixels. The auditory stimuli were presented via headphones. The participants were seated with their eyes approximately 60 cm from the display. Their head movements were unrestricted. A Tobii TX300 eye-tracker, sampling at 300 Hz, recorded their eye-movements while they listened to the experimental sentences. The participants received instructions asking them to listen to the sentences while looking at the display. On each experimental trial the participants’ task was to report whether the sentence they heard matched the visual images. For example, when the visual display contained pictures of a cook, a hippopotamus and a stick, but was accompanied by the auditorily presented sentence *The cook poked the elephant with the stick*, the participants were expected to answer “no”. The participants submitted ‘yes’ and ‘no’ answers by pressing designated keyboard keys. The auditory target sentences always matched the visual images. Half of the distractors did and the other half did not match the pictures.

A calibration was performed before the eye-tracking recording session. The participants first completed eight practice trials, and then proceeded to the main experimental block. On each trial, the participants first saw a fixation cross, shown for 1000 ms at the center of the screen. The visual stimuli were presented 500 ms prior to the onset of the auditory stimulus. The images remained on the screen for 300 ms after the offset of the sentences, and then were replaced by a verification question slide asking whether the pictures and the sentence matched. Next trial
started whenever participants submitted their response to the verification questions by pressing
designated key on a computer keyboard. The experiment lasted for approximately 20 to 25
minutes.

**Data processing**

On each trial, recording of the eye movements started at the onset and continued until the
offset of an auditory stimulus. The eye-tracking data comprised recordings of the participants’
fixations on the visual stimuli and were collected binocularly at the rate of 300 data points per
second. Each data point contained information about presence (or absence) of participants’ eye
gazes within the region of each image in the visual display. In order to examine anticipatory eye
movements, I extracted two temporal intervals from the eye-tracking recordings\(^6\). The first subset
contained fixation data for the period of the target locative verbs’ duration. The second fragment
included eye-tracking recordings for the auditory duration of direct objects comprising
participants’ fixations during both determiner and noun segments.

The raw fixation data for each ROI were aggregated into 50 ms bins. The aggregation
was done by participants and items. For each bin, I computed *empirical logit* (*elog*) value using
the formula in (8).

\[
(8) \quad elog = \log((y + .5)/(N - y + .5))
\]

In this formula, \(y\) stands for the number of times an event of interest (i.e., participant’s
fixations on a picture) is observed; \(N\) is a total number of observations, with \(N - y\),
correspondingly, standing for the number of “failures” to observe the target event. Empirical
logit is an approximation of a log transformed proportion of “successes” to “failures” (i.e., log of
odds) which safeguards from its positive/negative infinity values when the proportion of

\(^6\) Dink and Ferguson (2015) offer a comprehensive tutorial on how to handle eye-tracking data.
“successes” to “failures” equals 1 or 0, respectively (Jaeger, 2008). It does so by adding a 0.5 epsilon value to both numerator and denominator. The empirical logit, therefore, is the transformation of the categorical data into the log-likelihood metric. In the context of the current experiment, the elogs serve as the measure of likelihood of looking on a target image.

Similarly to Experiment 1, I discarded the experimental trials containing locative verbs unfamiliar to L2 speakers. Also, I removed the trials on which participants made errors in responding to the verification question: This affected 0.83% of the L1 English group data, 4.3% of the High L2; 5.76% of the Intermediate L2, and 7.5% of the Low L2 speakers’ recordings. Additional data reduction occurred due to a recording failure for two participants in the Low L2 group. The final dataset contained 99.17% of the L1 English data, 78.76% of High L2, 70.07% of Intermediate L2, and 64.50% of Low L2 participants’ eye-tracking recordings.

**Statistical Analysis**

The research question of Experiment 2 asked whether the study participants anticipated the upcoming verbal arguments of the non-alternating locative verbs. I examined two aspects of the oculomotor behavior relevant for an understanding the predictive processing. First, I investigated how participants’ eye fixations on the target images changed over the course of the verbs’ duration. In such analysis, if the participants do predict the upcoming arguments, the expected pattern of results would show an asymmetric increase in their looks to the respective targets and competitors. For instance, for the content non-alternating sentences such as *The cook spilled the coffee on the jeans*, we would expect a statistically significantly greater increase of looks to the image of *coffee* (in comparison to the picture of *jeans*) as the locative verb *spilled* unfolds. Correspondingly, the opposite pattern of fixations would be expected for the container-only sentences such as *The cook stained the jeans with coffee.*
The alternative analysis of the anticipatory sentence processing examined how likely the participants were to look at the content and container images right before the auditory presentation of the target nouns (Barr, 2008; Barr, Gann, & Pierce, 2011). Therefore, it investigated whether any anticipatory patterns observed in the time course analysis resulted in the target predictions post-verbally. I investigated the participants’ predictions by looking at the eye-tracking data associated with three temporal points in the auditory stream: (a) onset of the determiner (b) offset of the determiner, and (c) offset of the determiner plus 180 msec. I used these temporal locations in order to examine participants’ predictive behavior in its most inclusive terms. On the assumption that the predictions can be revealed later in the processing stream (for discussion, see Altmann, 2012; Altmann & Kamide, 2004; Huettig et al., 2011), the inspected points each provide additional information about the target behavior. At the onset of the determiner, nothing in the auditory stream reveals an identity of the upcoming argument. Therefore, witnessing participants’ preferences for the target images at this moment would serve as the clearest evidence of their predictive behavior. At the offset of the determiner, listeners have heard minimal information which is, however, quite unlikely to influence their eye fixations as an additional time (around 180-200 ms) is required to program the eye movements toward the target image (Huettig et al., 2011). Finally, the determiner offset plus 180 ms point is the latest moment in the auditory stream which still can be considered as being prior to the noun onset. At this moment, the oculomotor behavior might be influenced by the subtle auditory cues of the determiners, but not the nouns themselves.

I investigated both the trajectories of the anticipatory eye fixations over the course of the verb duration, as well the participants’ post-verbal predictions by using *Growth Curve Analysis*. The growth curve modeling is a statistical technique belonging to a family of multilevel linear
regression models (Mirman, 2014; Mirman, Dixon, & Magnuson, 2008). A defining feature of the growth curve modeling is that it incorporates time as a continuous predictor variable, making it possible to estimate changes in participants’ performance over time. An important aspect of this approach is that it can also accommodate non-linear trajectories. The non-linearity is typical of the “visual world paradigm” studies: In many cases, the eye movement patterns include rapid shifts in participants’ attention to a target or distractor as evident by accelerations/decelerations of the trajectories across a window of analysis. To estimate the curvilinear trends, the growth curve analysis incorporates higher order polynomial functions (e.g., quadratic, cubic, quartic, etc.) which capture the rises and falls in the data.

In running the growth curve analysis, an investigator has a choice between natural and orthogonal polynomials. One shortcoming in using natural polynomial transformations of time is that these predictors (e.g., $Time_{linear}$, $Time_{quadratic}$, $Time_{cubic}$, etc.) are highly correlated: the increase in a linear term necessarily coincides with an increase in a quadratic term, etc. Use of the orthogonal polynomials is a way of dealing with the problem of collinearity, which, however, changes interpretation of the growth curve models’ main effects at the intercepts. In the models with orthogonal polynomials, the main effects at the intercept signify differences between the conditions’ overall means (analogous to area under curve) rather than the differences between the experimental conditions at “Time 0”. Therefore, when the statistical analysis tests the differences at “Time 0”, the natural polynomials can be employed, otherwise, the orthogonal polynomial approach offers more sound means for assessing experimental manipulations.

To conduct the time course analysis, I used growth curve models examining the gaze data from onset to offset of the target locative verbs. I ran separate models for each experimental group estimating the likelihood (i.e. empirical logit) of looking at the target visual images. The
time course of fixations was modeled with a second-order polynomial and fixed effects of Verb Type (Content- vs. Container-only) and Image (Content vs. Container), as well as their interaction on all time terms. The random effect structure included by-subject and by-item random intercepts together with by-subject and by-item random slopes for all time terms (Barr, Levy, Scheepers, & Tily, 2013). To obtain p values for the predictor variables and their interactions, I compared models with and without the variables in question using the likelihood ratio test.

The investigation of participants’ post-verbal predictions involved similar statistical procedures. The eye-tracking data for three temporal intervals of interest starting from (a) onset of the determiner, (b) offset of the determiner, and (c) offset of the determiner plus 180 ms and then lasting till the offset of the direct object noun were examined using growth curve analysis. The models estimated the likelihood of participants’ eye fixations with Verb Type (Content- vs. Container-only) and Image (Content vs. Container) as fixed variables and included both participants and items in their random effect structure. The key difference from the previous analysis was the use of natural polynomials allowing the assessment of the effects of the predictor variable on the “Time 0” points.

**Results**

**Overview of the eye moment patterns**

Figures 11, 12, 13, and 14 illustrate the time course of participants’ fixations on the target images over the temporal interval stating from the onset of the locative verb and continuing until the offset of the target verbal argument. The vertical dotted lines indicate mean time of onsets of the determiner “the” and direct object noun for the content non-alternating (left panel) and container non-alternating (right panel) locative sentences. Figures 11, 12, 13, and 14 show the
fixation data for the L1 native, High L2, Intermediate L2, and Low L2 experimental groups, respectively.

Figure 11. Time course graph for L1 English group showing likelihood of fixations to the Content (black circles) and Container (white triangles) visual images viewed during auditory presentation of Content-only (left panel) and Container-only locative sentences.

Figure 12. Time course graph for High L2 English group showing likelihood of fixations to the Content (black circles) and Container (white triangles) visual images viewed during auditory presentation of Content-only (left panel) and Container-only locative sentences.
Figure 13. Time course graph for Intermediate L2 group showing likelihood of fixations to the Content (black circles) and Container (white triangles) visual images viewed during auditory presentation of Content-only (left panel) and Container-only locative sentences.

Figure 14. Time course graph for Low L2 group showing likelihood of fixations to the Content (black circles) and Container (white triangles) visual images viewed during auditory presentation of Content-only (left panel) and Container-only locative sentences.

There are two apparent visual patterns in the graphed data. First, there is a rapid rise of fixations to the images of content (e.g., wine) across all experimental groups beginning shortly after the onset of the locative verbs. The increase of looks to the content pictures is characteristic of both types of the locative verbs but appears to be more pronounced for the content non-alternating sentences (e.g., spilled the wine…). Second, a notable acceleration of fixations to the
images of containers (e.g., *t-shirt*) for the container non-alternating condition (e.g., *stained the t-shirt...*) emerges only by the offset of the locative verbs and then increases steadily during the determiner interval for all experimental groups except for low proficiency L2 learners. The following statistical analysis examined these preliminary observations.

**Time course analysis**

*English native speakers.* Figure 15 shows the average log-likelihood scores for participants’ fixations on content (blue triangles) and container (red dots) images viewed during auditory presentation of the verbal segment of content and container non-alternating locative sentences.

![Figure 15. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the verb segment of content-only (left panel) and container-only locative sentences for L1 English group.](image)

The growth curves are shown superimposed on the observed fixation data, and the full statistical results of the initial model are summarized in Table 21 in Appendix N. On average, over the course of verbs’ duration, native speakers of English looked more on the images of content (*\( \beta = 0.19, SE = 0.05, t = 4.00, \chi^2 (1) = 15.96, p < .0001 \)). The overall dominance of the
content images was characteristic of both categories of non-alternating locative verbs as suggested by the lack of statistically significant interaction between Image and Verb Type predictor variables ($\beta = 0.06$, $SE = 0.09$, $t = 0.65$, $\chi^2 (1) = 0.42$, $p = 0.52$). Also, there was a significant effect of Image on the linear term ($\beta = 1.46$, $SE = 0.23$, $t = 6.49$, $\chi^2 (1) = 42.08$, $p < .0001$) suggesting more rapid growth of looks towards the content images across the window of interest. At the same time, the statistically significant three-way interactions among Image, Verb Type and both linear and quadratic time terms suggested that the growth rate of fixations towards the target images differed between content and container non-alternating sentences (Image * Verb Type * Time linear, $\beta = 2.09$, $SE = 0.45$, $t = 4.62$, $\chi^2 (1) = 21.32$, $p < .0001$; Image* Verb Type *Time quadratic, $\beta = 1.44$, $SE = 0.42$, $t = 3.40$, $\chi^2 (1) = 11.58$, $p < .0001$). To examine the three-way interactions, I ran additional growth curve models separately for the content-only and container only locative sentences. The models’ results are summarized in Table 7.

Table 7. Summary of growth curve regression results for content and container non-alternating sentences for L1 English group

<table>
<thead>
<tr>
<th></th>
<th>Content Only</th>
<th>Container Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$</td>
<td>$SE$</td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.40</td>
<td>0.17</td>
</tr>
<tr>
<td>Image</td>
<td>0.22</td>
<td>0.07</td>
</tr>
<tr>
<td>Time linear</td>
<td>2.40</td>
<td>0.71</td>
</tr>
<tr>
<td>Time quadratic</td>
<td>0.76</td>
<td>0.34</td>
</tr>
<tr>
<td>Image * Time linear</td>
<td>2.56</td>
<td>0.33</td>
</tr>
<tr>
<td>Image * Time quadratic</td>
<td>0.91</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Notes. * $p < .05$, ** $p < .01$, *** $p < .001$, † $p = .07$
Native-speaking participants demonstrated greater growth of fixations to the content images (in comparison to the container image) during the processing of content non-alternating stimuli, as evidenced by the statistically significant effects of the Image predictor on both linear ($\beta = 2.56, SE = 0.33, \chi^2 (1) = 59.78, p < .0001$) and quadratic terms ($\beta = 0.91, SE = 0.30, \chi^2 (1) = 9.16, p < .0001$). The linear growth of fixations to the container and content images did not differ statistically for the container non-alternating sentences ($\beta = 0.47, SE = 0.30, \chi^2 (1) = 1.56, p = .12$). At the same time, there was a marginally significant effect of Image on the quadratic term ($\beta = -0.52, SE = 0.29, \chi^2 (1) = 3.27, p = .07$) indicating an acceleration of the fixation rate towards the images of containers in processing container-only verbs.

In sum, the analysis of time course showed an overall bias towards the content images for both types of locative verbs. The clear anticipatory fixation patterns were observed for the content non-alternating stimuli: Upon hearing such sentences as *The cook spilled the wine on the T-shirt*, native speakers of English tended to shift their looks towards the image of *wine* as the verb *spilled* unfolded. Even though such obvious anticipatory eye movements were not characteristic for the container non-alternants such as *The cook stained the T-shirt with wine*, the statistical results indicated that these stimuli did affect participants’ gaze patterns in the predicted way. Given that participants’ eye fixations for both categories of non-alternating verbs were recorded against the same set of visual images, the difference in the growth rates between two categories of locatives and, especially, the acceleration of fixations towards container images upon presentation of the container-only verbs suggest anticipatory effects of the container non-alternating locative verbs.
**High L2 group.** Figure 16 demonstrates mean observed values as well as growth curve model fits for the High L2 group’s eye fixations during auditory processing of the content non-alternating (left panel) and container (right panel) non-alternating locative sentences.

![Figure 16](image)

*Figure 16.* Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the verb segment of content-only (left panel) and container-only locative sentences for High L2 group.

Over the course of the locative verbs’ duration, High L2 speakers looked overall more on the images of content as evidenced by a statistically significant effect of Image on the intercept term ($\beta = 0.37, SE = 0.06, t = 5.91, \chi^2 (1) = 34.87, p < .0001$). The marginally significant three-way interaction among Image, Verb Type and quadratic term ($\beta = 1.09, SE = 0.57, t = 1.90, \chi^2 (1) = 3.63, p = .06$) suggested differences in growth rates of eye fixations to the visual stimuli for two categories of auditory stimuli (see Table 22 in Appendix N for a full summary of results). The results of the follow up analysis conducted separately for the content and container non-alternating sentences are summarized in Table 8. The rates of fixations to the content and container images did not differ statistically for the content non-alternating sentences, as demonstrated by a lack of statistically significant effects of the Image predictor on the linear ($\beta$
= 0.70, SE = 0.47, \( \chi^2 (1) = 2.18, p = .14 \) as well as on quadratic terms \( \beta = 0.41, SE = 0.43, \chi^2 (1) = 0.91, p = .34 \): The two curves increased in parallel (see Figure 16, left panel) with the overall preference for the content images across the inspected temporal interval \( \beta = 0.37, SE = 0.10, \chi^2 (1) = 14.83, p < .0001 \). The marginally significant effect of Image on quadratic term \( \beta = -0.68, SE = 0.38, \chi^2 (1) = 3.29, p = .07 \), observed for the container-only verbs, suggested that there was an increase in the rate of fixation to the container images.

Table 8. *Summary of growth curve regression results for content and container non-alternating sentences for High L2 group*

<table>
<thead>
<tr>
<th></th>
<th>Content Only</th>
<th>Container Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>( SE )</td>
<td>( \chi^2 (1) )</td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.34</td>
<td>0.16</td>
</tr>
<tr>
<td>Image</td>
<td>0.37</td>
<td>0.10</td>
</tr>
<tr>
<td>Time linear</td>
<td>1.62</td>
<td>0.59</td>
</tr>
<tr>
<td>Time quadratic</td>
<td>0.45</td>
<td>0.32</td>
</tr>
<tr>
<td>Image * Time linear</td>
<td>0.70</td>
<td>0.47</td>
</tr>
<tr>
<td>Image * Time quadratic</td>
<td>0.41</td>
<td>0.43</td>
</tr>
</tbody>
</table>

*Notes.* *p < .05,* ***p < .001,* †p = .07

Therefore, similar to native speakers, the High L2 participants showed a general bias towards the content images. Also, they demonstrated the theoretically expected anticipatory eye movement patterns in processing content-only stimuli, as well as a tendency to shift their visual attention towards the images of containers when hearing the container non-alternating locative verbs.

**Intermediate L2 group.** Figure 17 shows fixation data and growth curve fits for the Intermediate L2 group. The outcomes of the growth curve regression analysis (see Table 23 for the complete summary of results) indicated the overall preferences for the content images \( \beta = 0.46, SE = 0.07, t = 7.10, \chi^2 (1) = 50.34, p < .0001 \). The observed preferences, however, differed
statistically for content-only and container-only stimuli, as evidenced by a statistically significant interaction between Image and Verb Type independent variables ($\beta = 0.31$, $SE = 0.13$, $t = 2.34$, $\chi^2(1) = 5.64$, $p = .02$). Also, there was a statistically significant interaction among Image, Verb Type and linear time term ($\beta = 1.87$, $SE = 0.64$, $t = 2.90$, $\chi^2(1) = 8.43$, $p < .0001$), suggesting differences in growth of fixations to the content and container visual stimuli for content and container non-alternating sentences. Table 9 summarizes results of the follow up analysis ran separately for each category of non-alternating verbs.

*Figure 17.* Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the verb segment of content-only (left panel) and container-only locative sentences for Intermediate L2 group.
The regression results indicated that the participants in the Intermediate L2 group, overall, were more likely to look on the content images during processing of both content-only ($\beta = 0.63$, $SE = 0.10$, $\chi^2 (1) = 40.23$, $p < .0001$) and container-only locative sentences ($\beta = 0.32$, $SE = 0.08$, $\chi^2 (1) = 14.25$, $p < .0001$). Also, they demonstrated greater growth of fixations towards content images (in comparison to the container images) for both content non-alternating ($\beta = 2.97$, $SE = 0.49$, $\chi^2 (1) = 36.65$, $p < .0001$) as well as container non-alternating auditory stimuli ($\beta = 1.10$, $SE = 0.41$, $\chi^2 (1) = 7.20$, $p = .01$). Therefore, the statistically significant interactions observed in the initial model reflected more pronounced overall preference and greater difference in the growth rates for the content images during the auditory processing of the content non-alternating locative verbs. In both experimental conditions the participants showed the content bias which was significantly greater for the content-only sentences. The content bias was also indicated by the marginally significant effect of Image on the quadratic term for the content non-alternating verbs ($\beta = 0.83$, $SE = 0.44$, $\chi^2 (1) = 3.61$, $p = .06$) suggesting greater acceleration of looks towards the content image.
In sum, Intermediate L2 groups showed the general bias towards the content images in processing both types of the locative sentences. In contrast to the L1 and High L2 groups, the intermediate proficiency L2 speakers did not demonstrate the theoretically expected visual attention shift towards the container images in processing container-only verbs. At the same time, they did differentiate between the two types of the auditory stimuli: The differentiation, however, was reflected in the degree of the content bias for content and container non-alternating locative verbs.

**Low L2 group.** Figure 18 shows a summary of the fixation data as well as the growth curve model fits for Low L2 participants’ processing of content and container non-alternating verbs. The growth curve analysis results indicated that the low proficiency L2 speakers overall looked more on the content images as evidenced by a statistically significant effect of Image on the intercept term ($\beta = 0.38, SE = 0.07, t = 5.78, \chi^2 (1) =33.18, p <.0001$). Also, there was a statistically significant three-way interaction ($\beta = 3.01, SE = 0.64, t = 4.71, \chi^2 (1) =22.00, p <.0001$) suggesting differences in the growth rate of fixations on the target images for content and container non-alternating verbs.
Figure 18. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the verb segment of content-only (left panel) and container-only locative sentences for Low L2 group.

Table 10. Summary of growth curve regression results for content and container non-alternating sentences for Low L2 group

<table>
<thead>
<tr>
<th></th>
<th>Content Only</th>
<th></th>
<th>Container Only</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>SE</td>
<td>χ²(1)</td>
<td>β</td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.40</td>
<td>0.13</td>
<td></td>
<td>-1.45</td>
</tr>
<tr>
<td>Image</td>
<td>0.36</td>
<td>0.07</td>
<td>29.81***</td>
<td>0.31</td>
</tr>
<tr>
<td>Time linear</td>
<td>1.56</td>
<td>0.27</td>
<td>20.59***</td>
<td>1.35</td>
</tr>
<tr>
<td>Time quadratic</td>
<td>-0.01</td>
<td>0.16</td>
<td>0.01</td>
<td>-0.12</td>
</tr>
<tr>
<td>Image * Time linear</td>
<td>1.01</td>
<td>0.32</td>
<td>10.12**</td>
<td>-0.34</td>
</tr>
<tr>
<td>Image * Time quadratic</td>
<td>0.06</td>
<td>0.30</td>
<td>0.04</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Notes. ** p < .01, *** p < .001

The follow up analysis, summarized in Table 10, indicated that there was a statistically significant effect of Image on the linear term for content non-alternating verbs (β = 1.01, SE = 0.32, χ²(1) = 10.12, p < .01), but not for the container-non-alternating stimuli (β = -0.34, SE = 0.42, χ²(1) = 0.65, p = .42). These findings suggest that in processing the content non-alternating
sentences such as *The cook spilled the wine on the t-shirt*, the L2 speakers showed an asymmetric increase of eye fixations to the images of content (relative to the container images). This, however, was not the case for the container-only items such as *The cook stained the t-shirt with wine*: Processing the latter stimuli, participants demonstrated continuing preference for the content images with a parallel growth of fixations towards both content and container visual stimuli across the duration of the target verbs.

**Summary of the growth curve analysis**

The growth curve analysis revealed several notable patterns. First, all groups of participants generally looked more on the content images over the course of the stimulus duration: The content bias was observed across all proficiency levels and was characteristic of both categories of non-alternating locative verbs. Also, all experimental groups showed distinct patterns in the growth of fixations towards the target images for two categories of the locative non-alternants. For the content-only locatives, the participants demonstrated the theoretically expected continuing preferences and gradual increase of eye fixations on the content images. In contrast, for the container non-alternants, the content preferences and the increase in fixations were either significantly lower (for Intermediate and Low L2 speakers) or were accompanied by the shift of visual attention towards the container images (for High L2 and L1 speakers).

**Examining post-verbal predictions**

I examined the participants’ post-verbal predictions using growth curve models with 4th-order natural polynomial terms. I ran separate models for each group of participants. In each case, after running an initial analysis including both categorical predictors (Image = Content vs. Container; Verb Type = Content Only vs. Container Only), I conducted a follow up analysis separately for the content and container non-alternating items. In the following, I report the
statistics showing the effects of the fixed predictions and their interaction on the intercept, thus, focusing exclusively on the critical time points prior to the auditory presentation of the target arguments. First, I report the relevant information of the initial models, and then lay out the statistical results for each category of the locative verbs examining the participants’ predictions at the critical temporal points.

Figure 19. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the direct object segment of content-only (left panel) and container-only locative sentences for L1 English group.

**English native speakers.** Figure 19 illustrates L1 English speakers’ observed data and the growth curve model fits for the direct object segment of the content and container non-alternants. The time “0” point in Figure 19 indicates the article onset. The vertical dashed lines show mean offset of the article (170 ms) as well as the mean article offset + 180 ms critical regions. The initial growth curve models showed statistically significant interactions between Image and Verb Type for all temporal points of interest (article onset $\beta = 0.96, SE = 0.30, \chi^2 (1) = 10.38, p = .001$; article offset, $\beta = 2.27, SE = 0.31, \chi^2 (1) = 53.68, p < .0001$; article offset +180 ms, $\beta = 3.26, SE = 0.33, \chi^2 (1) = 99.84, p < .0001$) suggesting that the likelihood of participants’ fixations on the
content and container images depended on the verb type in the auditory stimuli. In processing the content-only sentences (e.g., *The cook spilled the wine on the t-shirt*), the participants were consistently more likely to look at the image of content (e.g., *wine*) before the target noun appeared in the auditory stream. The statistically significant results were found at the onset of the determiner ($\beta = 1.00, SE = 0.19, \chi^2 (1) = 27.55, p < .0001$), offset of the determiner ($\beta = 2.02, SE = 0.22, \chi^2 (1) = 87.27, p < .0001$), as well as at the determiner offset +180 ms temporal points ($\beta = 2.05, SE = 0.23, \chi^2 (1) = 79.65, p < .0001$). A different pattern was observed for the container-only locative sentences such as *The cook stained the t-shirt with coffee*. Neither at the article onset ($\beta = 0.12, SE = 0.21, \chi^2 (1) = 0.31, p = .58$), nor at the article offset $\beta = -0.24, SE = 0.22, \chi^2 (1) = 1.21, p = .27$), were there statistically significant differences between participants’ looks at the target images. Only at the article offset +180 ms temporal point native English speakers were more likely to look on the image of the container (e.g., *t-shirt*), therefore demonstrating the delayed predictions of the target locative arguments ($\beta = -1.20, SE = 0.23, \chi^2 (1) = 27.38, p < .0001$).

**High L2 group.** Figure 20 shows a summary of the High L2 group’s eye-tracking data with the growth curve model fits superimposed on the observed log-likelihood scores. The initial growth curve analysis indicated that at the onset of the determiner, the participants were more likely to look at the images of content irrespective of the verb type, as evidenced by the statistically significant main effect of Image ($\beta = 0.74, SE = 0.20, \chi^2 (1) = 14.26, p < .0001$), and a lack of the statistically significant interaction between the Image and Verb Type ($\beta = -0.03, SE = 0.39, \chi^2 (1) = 0.01, p = .95$). The differences between content and container non-alternating stimuli were observed at the determiner offset and determiner offset +180 ms points (article offset,
Image*Verb Type, $\beta = 1.13, SE = 0.40, \chi^2 (1) = 7.61, p = .006$; article offset+180, Image*Verb Type, $\beta = 0.98, SE = 0.43, \chi^2 (1) = 5.19, p = .02$).

Figure 20. Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the direct object segment of content-only (left panel) and container-only locative sentences for High L2 group.

The follow up analysis showed that for content-only stimuli the high proficiency L2 participants were more likely to look at the content images at the onset ($\beta = 0.72, SE = 0.29, \chi^2 (1) = 6.28, p = .01$) as well as offset ($\beta = 0.94, SE = 0.30, \chi^2 (1) = 9.66, p = .002$) of the determiner, but showed no preferences for either image kind on the delayed point of interest ($\beta = 0.39, SE = 0.32, \chi^2 (1) = 1.44, p = .23$). In processing the container non-alternating structures, they showed statistically significant preference for the content images at the onset of the article ($\beta = 0.75, SE = 0.26, \chi^2 (1) = 8.07, p = .004$), were equally likely to look at either content or container images at the determiner offset ($\beta = -0.19, SE = 0.27, \chi^2 (1) = 0.46, p = .50$), and then demonstrated more looks at the container visual stimuli at the article offset +180 ms point ($\beta = -0.60, SE = 0.29, \chi^2 (1) = 4.29, p = .04$).
In sum, the High L2 group’s fixation patterns can be best described as a crossover phenomenon. The content image preferences appeared early for the content-only locatives and then faded out at the delayed processing stage. The reverse pattern was characteristic of the container-only stimuli where the theoretically expected container preferences reached the statistically significant level only at the delayed processing point.

**Intermediate L2 group.** Figure 21 displays fixation data and model fits for Intermediate L2 speakers’ processing of the non-alternating locative sentences. The initial growth curve analysis revealed participants’ preferences for content images at the determiner onset.

![Graph showing fixation data and model fits for Intermediate L2 speakers' processing of non-alternating locative sentences.](image)

*Figure 21.* Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the direct object segment of content-only (left panel) and container-only locative sentences for Intermediate L2 group.

(Image, $\beta = 1.26, SE = 0.22, \chi^2 (1) = 33.11, p < .0001$; Image * Verb Type: $\beta = 0.04, SE = 0.44, \chi^2 (1) = 0.01, p = .92$), with the statistically significant differences between content and container...
non-alternating stimuli emerging at the offset of the determiner (Image * Verb Type, $\beta = 0.9$, $SE = 0.42$, $\chi^2 (1) = 4.87$, $p = .03$) as well as at the offset of the determiner + 180 ms points (Image * Verb Type, $\beta = 1.35$, $SE = 0.44$, $\chi^2 (1) = 9.22$, $p = .002$).

The results of the follow up analysis showed stable preferences for the content images in the processing of the content non-alternating locatives across all temporal points of interest (article onset, $\beta = 1.23$, $SE = 0.31$, $\chi^2 (1) = 16.01$, $p < .0001$; article offset, $\beta = 1.88$, $SE = 0.30$, $\chi^2 (1) = 38.99$, $p < .0001$; article offset +180 ms, $\beta = 1.29$, $SE = 0.13$, $\chi^2 (1) = 92.19$, $p < .0001$). In processing the container-only locatives, the Intermediate L2 speakers showed preferences for the content images at the onset ($\beta = 1.28$, $SE = 0.31$, $\chi^2 (1) = 17.43$, $p < .0001$) as well as at the offset ($\beta = 0.95$, $SE = 0.29$, $\chi^2 (1) = 10.52$, $p = .001$) of the determiner, and were equally likely to look at the content and container images at the determiner offset +180 ms point ($\beta = -0.04$, $SE = 0.30$, $\chi^2 (1) = 0.02$, $p = .90$).

To summarize, Intermediate L2 speakers demonstrated predictions for the target content arguments of the content non-alternating locative sentences such as *The cook spilled the wine on the t-shirt*: Prior to the onset of the auditory stimulus (i.e., *wine*) they consistently looked at the expected image. In processing the container-only sentences, such as *The cook stained the t-shirt with wine*, during the post-verbal interval, they demonstrated a shift from the generally preferred content images (i.e., *wine*) towards the theoretically expected images of containers (i.e., *t-shirt*). The observed shift, however, did not result in statistically significant preferences for the theoretically expected target argument.

**Low L2 group.** Figure 22 shows the eye-tracking data and the growth curve model fits for Low L2 speakers’ processing of the direct object segment of the target locative sentences. The initial growth curve models indicated consistent differences in the participants’ fixations to the
content and container images during auditory presentation of the content and container non-alternating sentences as evidenced by interactions between Image and Verb Type on each inspected point of interest (article onset, $\beta = 2.01, SE = 0.40, \chi^2 (1) = 25.07, p < .001$; article offset, $\beta = 1.96, SE = 0.42, \chi^2 (1) = 22.07, p < .0001$; article offset +180 ms, $\beta = 1.55, SE = .44, \chi^2 (1) = 12.48, p = .0004$).

![Figure 22](image-url)

**Figure 22.** Fixation likelihood for content (blue triangles) and container (red circles) images viewed during auditory presentation of the direct object segment of content-only (left panel) and container-only locative sentences for Low L2 group.

The follow up analysis revealed that the Low proficiency L2 speakers consistently preferred the images of content during processing of the content-only sentences (article onset, $\beta = 1.67, SE = 0.28, \chi^2 (1) = 33.87, p < .001$; article offset, $\beta = 1.68, SE = 0.30, \chi^2 (1) = 31.30, p < .0001$; article offset+180 ms, $\beta = 1.56, SE = 0.32, \chi^2 (1) = 24.31, p < .0001$), but were equally likely to look at the content and container visual stimuli for the container-only auditory stimuli.
Summary of the experimental results

The results of Experiment 2 revealed complex but consistent patterns in the participants’ processing of the target locative structures. Over the total course of the verbs’ duration, all groups of the study participants, regardless of their proficiency in English were more likely to look at the images of Content. This general tendency to prefer Content images was characteristic of the processing of both content- and container-only sentences: Irrespective of whether the participants heard such sentences as *The waiter poured*... or *The waiter filled*..., their eye-movements were predominantly directed to the images of Content (e.g., *water, tea*). At the same time, the growth curve analyses demonstrated consistent differences between the processing of the content- and container-only items reflected in the manner in which the participants’ eye gazes gradually changed across the temporal interval of interest. As evident from the statistically significant three-way interactions among Verb Type *Image and Linear and/or Quadratic time terms observed across all experimental groups, the dynamics of the participants’ growth of fixations (i.e. increase/decrease in the likelihood of looking at a particular picture) depended on the type of test item they heard. When presented with the content-only sentences (e.g. *The waiter poured/spilled*...), at each moment of the unfolding verbs (e.g., *poured, spilled*) the participants were increasingly more likely to look at the Content pictures (e.g., *water*) as compared to the growth of their looks to the images of Containers (e.g., *glass*). The statistically significant results of the linear growth were observed for the L1 group as well as for the Intermediate L2 and Low L2 participants. The linear growth dynamics of High L2 group did not differ statistically between Content and Container images: The participants’ eye fixations increased at comparable rates for
both Content and Container images yet favoring the Content pictures over the total course of the verbs’ duration.

In contrast, the time-course of processing the container non-alternating sentences (*The waiter filled/covered…*) revealed the equivalent linear growth of eye-fixations towards both Content and Container images for a majority of the experimental groups (all except the Intermediate L2 group). As the container-only verbs, such as *filled, covered,* and *stained* unfolded, the participants’ linear growth of fixations to the images of Contents (e.g., *water, tea*) and Containers (e.g., *glass, cup*) proceeded at a statistically comparable pace. Importantly, two experimental groups, native English speakers and L2 learners of high proficiency, also showed rapid acceleration of the eye-fixations to the Container images in processing the container-only sentences, as evidenced by the statistically significant effects of Image on the Quadratic time term. In sum, the analyses of the participants’ eye movements during their processing of the locative verbs revealed (a) overall dominance of the eye fixations on the Content images, as well as (b) differing processing dynamics distinct for the content- and container-only locative verbs.

The differences in the oculomotor dynamics during processing of the verbal temporal interval were further observed in the participants’ post-verbal predictions. All experimental groups were more likely to look at the Content images at the onset of the direct objects’ determiners during processing of the content-only sentences, therefore demonstrating clear anticipation of the upcoming locative arguments. At the same time, the participants’ post-verbal eye fixations during processing of the container non-alternating structures formed an implicational cline summarized in Table 11. The native speakers of English showed no statistically significant preferences for either Content or Container images at the determiner onset and determiner offset but were more likely to look at the images of Containers at the article
offset+180 ms temporal point. Therefore, despite the overall Content bias evident in the processing of the verbal segments, the L1 group showed delayed anticipation of the upcoming arguments. The High L2 group demonstrated residual Content preferences at the article onset, showed no preferences at the article offset and, similarly to the L1 group, was more likely look at the Container images at the article offset + 180ms. The residual Content preferences were evident in the post-verbal processing of the Intermediate L2 group on both the article onset and article offset points. At the article offset+180ms, this experimental group showed no preferences for either of the target images. Finally, the Low L2 participants were equally likely to look at either Content or Container images across all three post-verbal temporal points examined in the study.

Table 11. Odds ratio of looking at Container images in processing container non-alternating locative constructions by experimental groups

<table>
<thead>
<tr>
<th>Temporal points of Interest</th>
<th>Article onset</th>
<th>Article offset</th>
<th>Article offset+180ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 English</td>
<td>0.89</td>
<td>1.27</td>
<td>3.32</td>
</tr>
<tr>
<td>High L2</td>
<td>0.47</td>
<td>1.21</td>
<td>1.82</td>
</tr>
<tr>
<td>Intermediate L2</td>
<td>0.28</td>
<td>0.39</td>
<td>1.04</td>
</tr>
<tr>
<td>Low L2</td>
<td>1.40</td>
<td>1.32</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Notes, Colors of the cells represent participants’ prediction status: yellow = equal likelihood of looking at either Content or Container images; red = anticipation of the “wrong” arguments (i.e., Content), green = anticipation of the Container argument
In sum, the experimental findings indicate the prevalent initial tendency of the study participants to look at the images of entities set in motion by the actors of the locative events during auditory comprehension of the locative sentences. This Content bias facilitates anticipation of the Content arguments in the processing of the content non-alternating structures and is evident in the oculomotor behavior of all experimental groups. At the same time, such bias works against the anticipatory processing of the container non-alternating locative constructions. Even though all participants demonstrated performance distinguishing content- and container-only test items, only the L1 and High L2 proficiency groups showed delayed anticipation of the container-only arguments.
CHAPTER 4: DISCUSSION

The experiments reported in this project were aimed at obtaining detailed and unambiguous empirical evidence needed to resolve inconsistencies of the existing proposals concerning the learning of locative semi-regularities in an L2. Two experimental studies exploring L2 learners’ offline (Experiment 1) and online (Experiment 2) perception of locative constructions revealed complex but consistent patterns in participants’ performance. In the following discussion, I first summarize the empirical findings and use these results to reevaluate the validity of the alternative proposals. Then, I revise the processing-based account accommodating those findings that had not been predicted initially but were later discovered through experimentation. I present concluding remarks providing directions for future investigations.

Acceptability Judgements Experiment

Experiment 1 examined L2 learners’ acceptability judgments of locative constructions. The first research question asked whether L2 English learners of low, intermediate, and advanced proficiency demonstrate target-like performance in their ratings of the test items. To determine whether L2 participants’ performance meets the target language standards, I compared the L2 learners’ ratings to the acceptability scores of L1 participants (i.e., benchmark comparison), examined whether L2 participants within each proficiency band showed preferences for the well-formed over ill-formed locative sentences (i.e., preferential comparison) and were likely to accept grammatical/reject ungrammatical constructions (i.e., absolute comparison).

Overall, the analysis yielded mixed results revealing both target-like performance as well as systematic divergences from the assumed target language standards. The L2 English learners...
of all proficiency levels differed from the L1 speakers: They were less likely to accept grammatical and more likely to accept ungrammatical locative constructions. At the same time, all experimental groups showed preferences for the well-formed realizations of the non-alternating locative verbs: The preferential comparison revealed that L2 learners of all proficiency levels were more likely to rate as acceptable such sentences as The waiter poured water into the glass/The maid filled the cup with tea when compared to their acceptability ratings of ungrammatical *The waiter poured the glass with water/*The maid filled tea into the cup. Finally, the assessment of the L2 learners’ ratings in absolute terms showed that none of the L2 groups rejected ungrammatical locative constructions demonstrating the overextension tendencies. Importantly, the degrees to which the study participants tolerated the ungrammaticalities formed clear implicational scales across the proficiency levels and differed for content and container non-alternating items. As expected, L1 participants were highly unlikely to accept ill-formed content-only locative constructions such as *The waiter poured the glass with water. In turn, the High L2 and Intermediate L2 English learners were as likely to accept as reject and Low L2 participants accepted the ungrammatical items. The similar (although shifted) pattern emerged from the analysis of the ungrammatical container-only constructions: Whereas L1 English speakers were undetermined about the acceptability of such items as *The maid filled tea into the cup, all L2 groups were more likely to accept than reject such sentences.

Each comparison brings a unique perspective to the participants’ experimental performance. We know that L2 learners’ acceptability ratings differ in a predictable manner from L1 speakers’ scores, and that L2 learners both make distinctions between well- and ill-formed locative constructions, and, at the same time, do not reject the dispreferred ungrammaticalities.
Assuming that all analyses are accurate and equally contribute to the overall description of the L2 learners’ experimental performance, it is nevertheless incoherent to maintain that the learners simultaneously have and have not mastered the target structures. An important question at this point is what inferences we can draw about L2 learners’ mental grammar from such observations.

The least contentious and generally predictable are the observations that L2 English learners do not match L1 speakers in their acceptability ratings. In fact, a long-standing argument in the L2 literature known as the “comparative fallacy” (Bley-Vroman, 1983) warns against making any strong conclusions about the nature of L2 grammatical knowledge solely from such L1-L2 non-convergences, and, instead, encourages researchers to treat (and examine) L2 learners’ developing system as a linguistic system in its own right. Therefore, the most sensible way to interpret the observed non-isomorphies between natives and L2 English learners (revealed by the benchmark comparison) is to treat them as the signs of uncertainties in L2 learners’ responses. This reasoning leaves the reader with a pair of mutually opposing observations: the target-oriented performance in preferential terms and the non-target-like tolerance of the ungrammatical sentences revealed by the absolute comparison. A logically plausible way to accommodate both findings is to assume that the observed preferences do reveal something about participants’ grammars which makes them more likely to accept the well-formed sentences (relative to the ungrammatical ones), and that the overextensions (shown by the absolute comparison) are the artifacts of the experimental task. The validity of such an assumption, in the first place, can be argued for on purely theoretical grounds: It is simpler to maintain that the participants have the target-oriented knowledge but tend to underperform under specific conditions of the experimental task, rather than assume no such knowledge and try to
derive the target-oriented preferences from some unknown causes. Also, the gradience in the L2 participants’ non-rejections of the ill-formed sentences (revealed by the absolute comparison) suggests a systematic relationship between participants’ English proficiency and their acceptability ratings. Such a relationship is conceivable if we assume that the divergences from the target performance arise from the task specific demands: The more proficient the participants are the less likely they are to tolerate the ill-formed constructions judging them as unacceptable. On the other hand, if we consider the non-target-like judgements as indications of incompetence, it is unclear why the more proficient L2 learners would display greater resistance to the overextensions.

In sum, adopting the alternative perspectives on the single dataset grants a detailed outlook at the participants’ experimental performance. It reveals systematic patterns in the acceptability judgments both converging and non-converging with the assumed target language standards. A crucial point of the current discussion is that the hypothesized uncertainties in L2 learners’ ratings, their preferences for the grammatical items and overextensions of the dispreferred locatives are the complementary elements jointly describing the states of L2 learners’ grammar, and, thus, must be effectively accommodated by any theoretical account concerning the development of locative constructions in L2.

Even a preliminary reexamination of the alternative theoretical accounts in light of the experimental findings reveals that none of the competing proposals predicted all experimental outcomes. Therefore, it is sensible in the course of the reexamination to further ask whether it is possible to introduce adjustments to the existing accounts without undermining their foundational assumptions and internal logic, and, in such a way, to try to reassess the alternative proposals in their most accommodating terms.
The position advocated by Bley-Vroman and Joo (2001) and Joo (2003) hinges upon the hypothesis that adult L2 learners, in principle, do not have the means to acquire the specifics of the locative semi-regularities and cannot rely on their L1 in learning the target constructions. Two out of three comparisons (i.e., the benchmark and absolute) indeed revealed empirical evidence potentially supporting this position. However, the observed preferences for the well-formed over the ill-formed locative sentences demonstrated by all experimental groups call into question the validity of such proposal. After all, it is unclear how L2 participants across all proficiency levels would display such experimental performance, and, at the same time, be characterized as unable to make such a distinction. A possible way to accommodate the observed preferences, as suggested by Bley-Vroman and Joo (2001), is to assume that L2 learners, unlike L1 English speakers, learn to associate particular verbs with particular constructions in an “unprincipled way”, and therefore, show the target-oriented behavior relying on categorically different (from English native speakers) mental mechanisms. Although such an argument is plausible and difficult to refute without further details of what exactly is the “unprincipled way”, at least two aspects of the experimental findings challenge this suggestion. First of all, it is questionable to attribute the categorical differences on the basis of consistent preferences demonstrated across all proficiency levels. Even Low L2 participants who supposedly had minimal exposure to the target constructions were more likely to accept the well-formed items (relative to their acceptability of the ill-formed items). The observed target-oriented performance arising from minimal experiences with the locatives by no means implies a lack of “principled means” to acquire the distinctions among the locative verbs. Neither does it support the attribution of fundamentally different mental mechanisms to the L2 learners who similar to L1 English speakers showed consistent preferences for the grammatical items: In purely theoretical
terms, it is unsound to advance more complex explanations, such as existence of categorically different learning mechanisms, when a simpler one suffices. Also, the scalar patterns of overextensions discovered by the absolute comparison entail continuity between L1 and L2 participants’ experimental performance: By looking at the L1 group’s ratings we are able to predict the degrees of overextensions displayed by L2 learners of different proficiency levels. When L1 speakers do not accept the ungrammaticality of the content non-alternating locatives, L2 learners of high and intermediate proficiencies tend to be equally likely to accept and reject these items, and the beginning L2 learners are more likely to judge them as acceptable. In turn, when native English speakers remain undecided about the acceptability of the container-only ungrammatical items, all L2 groups are more likely to accept them. Such an implicational relation undermines the claims that L2 learners are categorically different from L1 participants and instead points in the opposite direction: Most likely, the overgeneralizations displayed by L2 and L1 participants reflect the same underlying mechanisms producing non-target-like behavior under specific experimental conditions. Therefore, even though the outcomes of both benchmark and absolute comparisons might be tentatively interpreted in support of Bley-Vroman and Joo (2001) and Joo’s (2003) position, the preferential evidence challenges its core claim and cannot be accounted for by the proposed theoretical adjustment.

The essential part of Juffs’s (1998) account is that L2 development of locative constructions depends on cross-linguistic correspondences formally accounted for in terms of lexical parameter settings. In the case of Chinese L1 speakers learning L2 English, it, first of all, predicts transfer effects: positive for content-only and negative for container-only locatives, and then hypotheses that the switch of the parameter settings should result in the target-oriented representations of L2 locatives. The only aspect of the experimental findings which might be
tentatively interpreted in support of the parametric account is the differential rate of
overextensions observed in the absolute terms for the content- and container-only locatives.
However, such evidence is problematic as not only did L2 learners demonstrate differences in
overextension of content and container non-alternating constructions. Contrary to the assumed
standards, the L1 group showed greater tolerance of the container-only ungrammaticalities. They
were undetermined about the acceptability of such sentences as *The maid filled tea into the cup.
Similarly challenging for Juffs’s (1996) proposal are the acceptability judgements of the
grammatical content- and container-only sentences by the beginning L2 learners. The acceptance
rates of the Low L2 group for the content-only items (e.g., The waiter poured water into the
glass) were slightly lower ($M = 84.38\%$, $SD = 36.42$) than their acceptance rates of the container
non-alternating sentences such as The maid filled the cup with tea ($M = 94.41\%$, $SD = 23.04$). It
is highly unlikely that the acceptability scores for the content non-alternating constructions
signify positive transfer effects and the scores for the container-only sentences reveal the effects
of negative transfer. If that were the case, we would expect much greater disfavor of the
container-only constructions from L1 Chinese speakers beginning to learn L2 English whose
native language, in principle, does allow incorporation of the caused changes. These
observations, in effect, significantly weaken Juffs’s (1996) account, and require an additional (or
alternative) explanation for why the content- and container-only structures are treated differently
by not only L2 speakers but also L1 experimental participants.

The outcomes of the benchmark and preferential comparisons similarly do not conform to
Juffs’s (1996) thesis. The analysis comparing L2 learners’ acceptability judgments to L1
speakers’ ratings (the analysis exclusively used in Juffs (1996)) showed consistent differences
between L2 learners of all proficiency levels and natives for both content and container non-
alternating items. Also, all groups of L2 learners, irrespective of their English proficiency, preferred well- over the ill-formed realizations of both categories of locative verbs. Therefore, neither of the findings can be accounted for by Juffs’s (1996) proposal whose key assumption hinges on the asymmetry between content- and container-only locatives arising from the cross-linguistic mismatches between learners’ L1 and the target language.

A plausible solution potentially solving the disparity between Juffs’s (1996) theoretical proposal and the experimental findings (seemingly adopted by the author, for details see Juffs (1996) Discussion section) is to assume that even minimal exposure to the target language is sufficient for the parameter resetting, and that overextensions and uncertainties in L2 learners’ ratings reflect the complexity of constraining the overgeneral argument realization patterns, as well as performance errors. Juffs (1996) especially ascribes the difficulty in restraining the overgeneralizations to the continuing L1 influence suggesting that the putative translation equivalents of English container-only locatives (e.g., to fill, to decorate, and to cover) represent a wider grammatical pattern (i.e., allow both content and container arguments as their direct objects), and that narrowing down the overgeneralizations (unlike adding new representational elements) might be difficult even for advanced L2 learners. Such revisions significantly improve the fit between the experimental findings and Juffs’s (1996) theoretical proposal. At the same time, they undermine the explanatory adequacy of the postulated lexical parameter: It is unclear why we would need to assume the existence of the lexical parameter whose effects are undetectable even at an early stage of L2 development. Also, the revised account apparently sets aside the core of the issue it is designed to solve: the (non)retreat from overgeneralizations of content and container non-alternating locative verbs in the L2. It hypothesizes the continuing effects of L1 even though the overextensions are not limited to the container non-alternating
constructions (e.g., *The maid filled tea into the cup) but are also evident in the ratings (assessed in the absolute terms) of the content-only ungrammatical items (e.g., *The waiter poured the glass with water). Equally unrevealing are the references to performance errors as the possible explanations of the non-target-like L2 learners’ ratings. Juffs’s (1996) proposal is a property-theoretical account (i.e., a competence theory) hypothesizing about specific states of L2 grammar at different points in L2 development. By its very nature, it has little to say about real-time language use and correspondingly does not predict any specific deviations from the declared L2 grammar states, at most, allowing for some marginal and superfluous performance incidents related to limitations of our memory and attention. Without further elaboration, the performance errors are uninformative and thus can be merely disregarded. The findings of the current study cast doubt on the disposability of “performance errors.” It is highly unlikely that L2 participants’ distractions and memory limitations might result in the predictable and consistent pattern of uncertainties and overextensions revealed by the benchmark and absolute comparisons. In sum, Juffs’s (1996) proposal, in its original formulation, is not supported by the experimental findings reported in the current project. Furthermore, the attempts to introduce adjustments improving the theory-data fit undermine its explanatory adequacy considerably weakening its predictive capacities.

The core assumption of the processing-based account advanced in the current project is that both semantic interpretations, and, by extension, both structural realizations of locative arguments are available for all locative verbs. It is regarded as a default option arising from the inherent ambiguity of all locative events. I envisioned the development of the content- /container non-alternating constructions as a gradual increase in mnemonic accessibility of the syntax-
semantic correspondences involving statistical and semantic learning mechanisms embedded within the MOGUL framework.

Similar to the previously discussed proposals, the processing-based account failed to anticipate the patterns in the participants’ acceptability ratings, accounting only for a small fraction of the empirical findings: The Low L2 participants did differ from the L1 speakers and were more likely to accept ungrammatical content- and container-only locative constructions. However, contrary to the initial hypothesis, they also distinguished between well- and ill-formed test items being more likely to accept the grammatical constructions. Also, the L2 learners on the other end of the proficiency spectrum did not match the natives in their acceptability ratings and failed to fully retreat from overgeneralizing both content- and container-only ungrammaticalities.

Despite being unpredictable in its initial formulation, the processing-based account is readily amendable with a minimal set of adjustments. First of all, even though both construals of locative events are available for all locative verbs, minimal exposure to the target constructions (i.e., necessary for recognizing a verb as a lexical item of the target language) can be sufficient to make attested argument realization patterns more accessible in comparison to their competing alternatives. In MOGUL’s terms, each time L2 learners encounter locative verbs in one (but not the other) sentence frame, this representational chain acquires a higher resting activation level becoming a stronger competitor for future processing. Consequently, the relative inequality between the competitors results in greater chances for the witnessed (i.e., grammatical) constructions to be perceived as acceptable even at the beginning stages of L2 learning. Such relative dominance of the attested argument realization patterns over their alternatives by no means implies that the unattested realizations are fully eliminated from language learners’ grammar. Rather, under favorable conditions, the relatively dispreferred representational chains
can gain sufficiently high activation levels and be used in language processing producing ungrammatical linguistic structures. When L2 learners are presented with ungrammatical locative constructions such as *The waiter poured the glass with water or *The maid filled tea into the cup, they tend to interpret these sentences/fit the non-alternating verbs into the dispreferred constructions despite having the better means for expressing the non-alternating verbs. The acceptance of ill-formed constructions declines with proficiency: as the well-formed structures gain greater dominance, the unattested constructions become less accessible and thus more dispreferred. This pattern is evident in the participants’ ratings of content-only ungrammatical constructions: whereas the Low L2 group readily accepts such structures as *The waiter poured the glass with water, L1 speakers reject and Intermediate L2 and High L2 groups remain indecisive about their acceptability. However, for L2 learners, such a trend does not reach its extremes. Even High L2 learners do not attain total dominance of the well-formed (attested) over the ill-formed (unwitnessed) constructions. This potentially explains their indecisiveness about the acceptability of the content-only ill-formed sentences as well as uncertainties observed by in the benchmark comparison: In contrast to L1 speakers, who are likely to establish maximum activational dominance of well- over ill-formed representations, L2 learners are consistently less likely to accept the grammatical and more likely to accept the ungrammatical locative structures.

One pattern that does fit into the overall picture is the content-oriented bias observed in participants’ scores assessed by the absolute comparison. As discussed above, it is evident in the shifted scales of L2 learners’ non-rejections of the ungrammatical constructions. Such asymmetry cannot be explained solely in terms of L1 transfer, and, therefore, must be integrated as a general principle into the current account with a minimal set of additional
modifications. At this point in the discussion, my basic assumption is that there must be something special about the caused-motion senses of the locative events, and, correspondingly, the content-oriented sentence frames which results in the observed asymmetry. However, I will postpone a full-fledged examination of such a phenomenon and address it later in light of the complete set of empirical findings gathered by the current project. To sum up, with minimal adjustments to its original formulation and a temporary hold on a detailed explanation of the content-bias, the processing-based proposal offers an accurate description of the experimental findings advancing a plausible account of how locative semi-regularities might be acquired in L2 acquisition.

**Interim summary**

Examination of the basic research questions of whether L2 learners of different proficiency levels demonstrate target-like performance on the acceptability judgment task revealed complex but consistent patterns in the participants’ experimental behavior. The findings suggest that L1 Chinese speakers learning L2 English show both target-like performance and diverge from the assumed target language standards. I used the observed patterns to reassess the validity of the conflicting positions on the development of locative constructions in L2. Each theoretical proposal accounted for only a fraction of the empirical findings generally failing to explain all experimental outcomes. Therefore, in the course of reexamining the theoretical alternatives, I also considered a range of plausible amendments to these positions and assessed them in the most extended and accommodating terms. Neither Bley-Vroman and Joo (2001) and Joo (2003), nor Juffs (1996) could account for the observed patterns. In contrast, a modified version of the processing-based proposal provided an accurate descriptive framework accommodating the empirical outcomes observed in the participants’ acceptability judgments.
Developmental trends in L2 learners’ acceptability judgments

The second research question inquired as to whether L2 learners’ ratings of non-alternating locative constructions varied across proficiency levels. It aimed at discovering developmental trends in the participants’ judgements which, in turn, could be used to further reexamine the competing theoretical proposals. I used participants’ preference scores (i.e., differences between acceptability scores for grammatical and ungrammatical content-/container-only constructions) as the outcome variable and examined whether they differed across the L2 groups of Low, Intermediate and High proficiency levels. The results of the statistical analysis revealed a clear developmental pattern in the participants’ ratings: the higher proficiency L2 learners tended to make greater distinctions between well- and ill-formed locative constructions. However, such preferences were more pronounced and reached the statistically significant level only for the content non-alternating locatives. This finding cannot be accommodated by Bley-Vroman and Joo (2001) and Joo’s (2003) account whose central assumption is antithetical to the observed pattern. Similarly problematic is the developmental trend for Juffs’s (1996) proposal: Contrary to the parametric account, the growth of preferences for the grammatical items was primarily associated with content rather than container non-alternating constructions. In turn, the gradual increase of L2 learners’ preferences for the grammatical realizations of locative verbs is fully compatible with the revised processing-based account. Assuming that the non-significance of the between-group differences for container-only locatives is an additional manifestation of the content-bias, the gradual growth of the preference scores for the content non-alternating structures is what we would expect if the development of non-alternating locatives involved increase in the mnemonic accessibility of the witnessed structures, and, correspondingly, decrease of the resting activation level of their unattested
competitors. In sum, the observed trends in the acceptability judgments help to refine our understanding of the locative phenomena in L2 adding key details to the overall canvas of empirical findings.

**Frequency effects in L2 learners’ acceptability judgments**

The third research question of Experiment 1 explored the relationship between the frequency of non-alternating locative verbs and L2 learners’ preference scores for the experimental constructions containing these lexical items. Even though the experimental outcomes revealed a strong predictive association between lexical frequency and L2 learners’ preferences, this relationship was observed only for the High L2 group’s ratings of content non-alternating locatives. How can we explain such circumscribed frequency effects and what do they imply about L2 development of locative constructions?

Similar to Ambridge and Brandt (2013), Robenalt and Goldberg (2016) suggested that such sort of outcomes indicate that L2 learners, unlike native speakers, have reduced ability to use statistical mechanisms across most of their developmental span and might rely on frequency information only at the advanced level of L2 proficiency. Such a claim primarily stems from the assumption of Construction Grammar which views inferences about ungrammaticality of unattested constructions as statistically driven inhibitory effects. From this perspective, if L2 learners do not judge as less acceptable ungrammatical constructions containing high frequency verbs, they probably do not rely on the error-driven learning mechanism suppressing consistently losing (i.e., unattested) competitors. In my view, such a conclusion is too radical or at least premature.

First, the insensitivity to the statistical properties is not something exclusively associated with L2 learners, but, most likely, indicates a developmental phase in either L1 or L2 acquisition.
For example, in a similar experimental study investigating effects of semantic and statistical factors in L1 learning of locative constructions, Ambridge et al. (2012) showed that the youngest group of native English speakers (age 5-6), unlike older children (age 9-10) or adult participants, did not depend on frequency information in the predicted manner: For this group, the acceptability judgment scores for ill-formed locative sentences were not associated with the frequency of the target verbs, suggesting that the lack of the expected statistical influences may be an integral part of normal language development rather a sign of language learners’ deficiencies. Also, the statistical effects are likely to work in a different way envisioned by proponents of Construction Grammar. In Ambridge et al. (2012), the same group of young L1 learners did show statistically significant frequency effects when instead of the raw acceptability ratings the researchers used participants’ preference scores. The discrepancy between the two statistical tests suggests that the frequency of locative verbs, in the first place, was associated with a greater acceptability of grammatical constructions (i.e., rating values included in the computation of the preference scores), rather than a rejection rate of ill-formed locatives. Importantly, the same frequency effects were found for adult L2 learners in Ambridge and Brandt (2013) whose participants tended to judge as more acceptable well-formed constructions containing high frequency verbs but showed no statistical effects in their judgements of ill-formed items. What these findings suggest is that L1 and L2 learners are likely to follow the same or similar developmental paths, and do not differ categorically from each other. Moreover, both populations are likely to use frequency information in their judgements of well-formed constructions before they are able to do so for ungrammatical items.

It is also important to draw distinctions between statistical learning mechanisms hypothetically involved in the growth of locative constructions and statistical/frequency effects
detectable by empirical investigations. Whereas observations of frequency effects in experimental studies do imply something about the underlying statistical learning mechanisms, the absence of observable effects does not necessarily mean that such mechanisms are deficient or nonexistent. The particulars of experimental designs as well as specific tasks used in empirical studies might mask, either partially or fully, the effects of the underlyingly functioning processes. The outcomes of the current project prove this point. The advanced L2 participants in the current study were more likely to judge as unacceptable ungrammatical content-only sentences with high frequency verbs. Everything else being equal, this finding can be interpreted as supporting the error-driven statistical learning mechanism advocated by proponents of Construction Grammar. However, such inference leaves unexplained why advanced L2 English learners in Ambridge and Brandt (2013) did not perform in a similar manner, but instead showed effects of verbal frequency exclusively in their ratings of well-formed constructions. A plausible resolution of such inconsistencies can be found in the designs of two experiments. The current project tested only a subset of locative verbs (six out of 20 content- and container-only items) used by Ambridge and Brandt (2013) selecting only those verbs that were likely to be familiar to the beginning L2 participants. The selection criterion affected not only the quantity of the tested verbs but also the distribution of lexical frequencies within the target set of locatives. The “low” frequency locative verbs investigated by the current project could have been actually “high” frequency items in Ambridge and Brandt (2013). Such shifted distributional characteristics can potentially explain the differences in the experimental outcomes. In Ambridge and Brandt (2013), who examined a broader spectrum of the locative verbs, the lexical frequencies were associated with ratings of the well-formed structures, the primary targets of the statistical learning mechanism advocated by the current processing-based account. Provided that
strengthening of the witnessed (i.e., grammatical) constructions affects unattested (i.e., ungrammatical) items in relative rather than strictly proportional (i.e., through inhibition) terms, the participants ratings of the ungrammatical locatives could have varied enough to conceal any potential frequency effects. On the other hand, the findings of the current experiment allow an alternative but complementary view on the underlying statistical learning mechanisms revealing a different set of frequency effects while masking others. The observed null-effect of the statistical predictor on L2 learners’ judgements of well-formed constructions most likely reflects a ceiling effect. Given the relatively high frequency of the target locative verbs, the advanced L2 participants did not vary in their responses, consistently ranking the test items as highly acceptable. In turn, the uniformity in the acceptability judgements resulted in the inability of the statistical model to detect any relationship between the distribution of lexical frequencies and L2 learners’ judgments. At the same time, statistical effects emerged for the ungrammatical items whose relative in(accessibility) is predicted to depend on the activation level (relative accessibility) of their grammatical competitors. Therefore, the most sensible way to explain the differential frequency effects observed across different experiments is to assume a statistical learning mechanism relying on the relative mnemonical accessibility of competing mental structures rather than the error-driven suppression.

One obvious corollary of the current discussion is that L2 learners must reach a certain threshold in their L2 development in order to demonstrate any frequency effects on the acceptability judgment task. Also, there are multiple factors that might affect detectability of statistical influences in a population of L2 learners. A central premise of the acquisition-by-processing statistical mechanism is that statistical learning depends on how often a representation is used in processing (i.e., internal frequency) (and/or how readily it is available) rather than the
objectively measured occurrences of linguistic items in the environment (i.e., external frequency). Even though external frequency information (i.e., corpus frequencies) might accurately approximate the internal statistics of a fluent language user, such estimates are likely to be less than optimal predictors of the internal statistics in developing L2 learners. First, L2 learners, especially in formal language learning settings, might be exposed to subsets of L2 which inaccurately represent distributional properties of the target language. Second, even when L2 learners are exposed to linguistic samples veridically representing target language statistics, there might be a few direct correspondences between external frequency estimates and L2 learners’ internal counts: The moderately frequent items (in external terms) might be unknown or vaguely recognized by L2 learners and high frequency items might turn out to be of low frequency in internal terms. In general, the activation states of L2 learners’ representations (i.e., their internal statistics) are likely to be indirectly and imprecisely related to the objective statistical estimates of the language they learn. Also, a key property of L2 learners is that they already possess a fully functioning linguistic system which can affect the internal statistics of L2 mental representations. Some internally weak items might have high frequency counts on external measures and simply are not taken into account in the construction of representations (such as English articles for L1 speakers of article-less languages), or, otherwise, be rather sparse in the environment but gain high levels of activation due to their central role in L1 processing. Similarly, internally generated biases unrelated to L2 learners’ L1 might skew the mnemonic accessibility of certain linguistic constructions irrespective of their external statistics. A vivid example of such internally represented imbalance is evident in the content-oriented bias observed in the current project. Despite being equal in their externally computed frequencies, content and container non-alternating verbs and their constructions were treated differently by participants of
the current study. In the examination of the frequency effects, the content-only sentences judged by High L2 learners were the only test items reaching the assumed threshold, and, thus, revealing the strong association between external frequency estimates and acceptability ratings. Furthermore, a number of other factors specific to the L2 population but unrelated to the discrepancies between internal and external statistics might play a role in masking frequency effects. For instance, L2 participants might tend to avoid extreme values on the rating scale strategically “playing it safe” in their acceptability judgements, and, therefore, concealing the genuine effects of lexical frequencies.

To sum up, there are many ways in which frequency effects might be masked in empirical investigations of L2 learners. Such null results should not be taken as evidence of deficiencies in statistical learning mechanisms, but rather carefully reexamined through all available theoretical and empirical means needed to reveal the intricate character of statistical associations between environmental evidence and the developmental course of the L2 system. More telling for the current discussion are the frequency effects observed in High L2 learners’ ratings of the content-only constructions. Even though these findings are confined to a single category of locatives and require L2 learners to reach a certain threshold level, they make a strong case for an involvement of statistical mechanisms in L2 acquisition of locative constructions.

Summary

The examination of the fine-grained nuances of Experiment’s 1 empirical findings provide us with a detailed outlook at the L2 participants’ performance on the acceptability judgment task. Despite an apparent complexity, the empirical outcomes reveal clear and
discernable behavioral patterns explainable in terms of the processing-based account advanced in the current project.

**Processing locative constructions: Examining eye-tracking correlates**

Experiment 2 examined L2 learners’ real-time processing of locative constructions using visual world paradigm eye-tracking methodology. The research question of Experiment 2 asked whether L2 learners of different proficiency levels could anticipate the arguments of non-alternating locative verbs. I analyzed two interrelated aspects of their oculomotor behavior: (a) changes in the participants’ eye fixations associated with the auditory processing of the locative verbs, and (b) likelihood of looking at the “right kind” of the visual images post-verbally. On my initial assumption, if the study participants do predict the upcoming nouns, they should demonstrate gradual increases in their eye fixations on the appropriate visual images (i.e., Content for content-only and Container for container-only constructions) during auditory processing of the locative verbs, and, as a result, be more likely to look at these pictures at the post-verbal time points preceding auditory presentation of the direct objects.

The outcomes of the statistical analyses revealed less clear-cut but nevertheless highly informative patterns in the participants’ oculomotor behavior. All groups taking part in the study were more likely to look at the images of Content over the course of the verbal duration. This was true not only for the content non-alternating constructions such as *The waiter poured…*, but also for the container non-alternating stimuli (e.g., *The waiter filled ...*). At the same time, all groups showed differential growth rates of eye-fixations during the processing of two categories of locative non-alternants revealing distinct effects of the content- and container-only verbs on the participants’ oculomotor behavior. During auditory processing of the content-only verbs (e.g., *to pour, to spill, and to drip*), L1 English speakers and L2 learners of all proficiency levels
demonstrated the hypothesized anticipatory performance: as the verbs unfolded, they were becoming increasingly more likely to focus on the images of Content (e.g., water, tea, and juice), and, as a result, were more likely to look at these pictures (i.e., predict them) post-verbally. Such a theoretically assumed pattern did not bear out for the container non-alternating locatives. During the processing of the container-only verbs (e.g., to fill, to decorate, and to stain), the images of Content maintained their dominant status in the visual attention of all experimental groups. The eye-fixations on the Content and Container pictures increased in parallel for the L1, High L2, and Low L2 groups or showed statistically greater content-oriented growth in the performance of the Intermediate L2 participants. Importantly, L1 English speakers and High L2 learners also displayed a growth trend running against the dominant (i.e., content-oriented) current: Both groups showed statistically detectable accelerations in their container-oriented growth rates suggesting a rapid shift of their visual attention towards Container images. For the L1 participants, this attentional shift produced the following post-verbal outcomes. At the onset as well as offset of the direct objects’ article, the native English speakers were equally likely to look at either Content or Container images, and at the latest prenominal time point (i.e., article +180 ms) they were more likely to look at the pictures of Containers. The similar eye movement dynamics were observed for the advanced L2 learners. They were more likely to focus on the Content images at the article onset, displayed no preference for either of the alternative pictures at the offset of the article, and, then, demonstrated significantly greater likelihood of looking at the Containers at the “delayed prediction” time point (i.e., article +180 ms).

In turn, the oculomotor performance of Intermediate L2 and Low L2 groups revealed similar post-verbal trends, which, nevertheless, did not result in the Container preferences even at the delayed temporal point of interest. At the onset as well as offset of the direct objects’
article, the Intermediate L2 participants were more likely to focus on the images of Content, and then displayed no preferences for the visual alternatives at the article offset +180ms time point. As for the Low L2 participants, this group was equally likely to look at either Content or Container pictures across all three post-verbal time points.

In sum, the findings of Experiment 2 revealed highly consistent patterns in the experimental performance of L1 and L2 English speakers. All participants displayed the strong content-oriented bias during processing of the locative constructions. Similar to the observations of Experiment 1, this bias differentially affected participants’ perception of the content and container non-alternating sentences. In processing the content-only (i.e., pour-kind) constructions, all experimental groups were able to anticipate the upcoming verbal arguments. On the other hand, such bias significantly hampered participants’ predictive performance in processing container-only items. When listening to such sentences as The waiter filled the …, all experimental groups were dealing with the overwhelming content dominance (especially on the earlier phases of auditory processing), and, as a result, only L1 speakers and High L2 learners could overcome the content-oriented influence switching their visual attention to the appropriate images at the delayed time points. In contrast, the Intermediate L2 and Low L2 groups, despite demonstrating a similar kind of attentional shift, did not fully recover from the content-oriented bias and were equally likely to focus on either visual image post-verbally.

An important step towards a better understanding of the current experimental outcomes is to compare them to the observations of Experiment 1. On the acceptability judgment task in Experiment 1, the participants responded to two structural realizations of each locative verb providing the rating scores used in the absolute, benchmark, and preferential comparisons. In contrast, the online task of Experiment 2 can be primarily considered as an alternative
measurement of the participants’ preferences examined within a single experimental trial: It relies on listeners’ tendencies to choose (or prioritize) among the visually perceived alternatives those entities that are mentioned or anticipated to be mentioned in the unfolding linguistic stream. Also, the eye-movement patterns allow applied linguists to infer how likely a group of language users would be to accept/reject a particular kind of locative construction, as such, constituting a processing analogue of the absolute analysis used in Experiment 1. A strong and persistent preference for one category of the visual targets over their competitors is likely to indicate acceptance of the preferred interpretations and correspondingly rejection of the ignored alternatives. On the other hand, unstable and shifting patterns of eye-fixations can be best understood as the signs of continuing competition associated with greater chances of accepting ill-formed test items.

Having established the tentative correspondences between the measurements of two experiments, it is informative to compare the acceptability and processing findings. When listening to the content-only locative sentences all participants demonstrated stable preferences for the content images uniformly predicting their appearance in the upcoming speech stream. Such an observation mirrors the outcomes of the preferential analysis of Experiment 1 revealing reliable preferences for the well-formed (e.g., *The waiter poured water into the glass*) over the ill-formed (e.g., *The waiter poured the glass with water*) content-only locatives demonstrated by L1 and L2 English speakers of all proficiency levels. At the same time, such an eye-movement patterning suggests that the L1 and L2 participants are not naturally inclined to attribute unintended semantic interpretations to the content-only verbs, and, under normal circumstances, they would be highly unlikely to perceive as grammatical the ungrammatical locatives. What this implies is that non-rejections of the ungrammatical items (observed in the
absolute comparison) are primarily the artifacts of the test conditions of Experiment 1. As suggested previously, when presented with the ill-formed sentences such as *Bryan spilled the t-shirt with coffee, L2 learners tended to make do with the provided test items interpreting them as grammatical despite having better/preferred means to express the locative events.

A somewhat different picture emerges when one compares the findings of the two experiments for the container non-alternating locatives. Even though all experimental groups demonstrated consistent preferences for the well-formed (e.g., Mary stained the jeans with wine) over the ill-formed (e.g., *Mary stained wine onto the jeans) constructions in the preferential analysis of acceptability judgments, the eye-tracking findings do not fully corroborate these observations. Only native English speakers and advanced L2 learners were found to predict (prefer) the upcoming Container arguments and did so only at the delayed temporal stages of auditory processing, whereas Intermediate L2 and Low L2 English learners were equally likely to focus on either visual image prior to hearing direct object stimuli. How can we resolve such interexperimental inconsistencies? A reasonable way to deal with such mismatches is to assume that the experimental conditions of the eye-tracking investigation posed more stringent requirements on the participants’ real-time performance, amplifying the existing content-oriented biases. The primary evidence in support of this assumption is the gradient predictive performance of the study participants in processing container-only constructions: The post-verbal predictions form a clear implicational cline structured along the English proficiency level. Provided that L2 leaners are generally slower to process linguistic information but gradually improve across the L2 developmental span (Hahne & Friederici, 2001; Lew-Williams & Fernald, 2010), it is likely that L2 learners of the lower proficiencies similar to the more fluent English users did anticipate/prefer the container arguments but were incapable of demonstrating these
preferences due to less efficient processing capacities. The processing speed, however, cannot fully account for the cross-study inconsistencies, but should be considered as an important factor in participants’ recovery from the content-oriented bias. None of the experimental groups demonstrated unstable and shifting patterns in their oculomotor behavior, but rather tended to prioritize content images in the earlier processing stages and then switched towards the container targets. It is possible that the visual displays themselves served as the semantically interpretable stimuli and could affect the eye-fixation patterns in parallel with the linguistically delivered interpretations. After all, the visual world paradigm investigations are about integration of visual and linguistic information, and, correspondingly, are not necessarily pure measures of how linguistic stimuli are processed but can also reflect influences of visual perception on the interpretation of language (Salverda, Brown, & Tanenhaus, 2012; Tanenhaus & Brown-Schmidt, 2008). Given the content-oriented bias, when the study participants saw the visual displays, they could interpret these visual scenes as depictions of caused-motion events even before they could fully integrate the critical locative verbs, and, as a result, demonstrated the eye-movement patterns reflecting complex interactions between language and vision.

Importantly, even though the specific conditions of the visual world paradigm might have affected the experimental observations, they by no means are the primary source of the content-bias. What this means is that even if all groups of L2 learners are likely to expect/prefer the right kind of arguments for the container-only locatives (i.e., show experimental performances matching the preferential findings of Experiment 1), they are more likely to be affected by the content-oriented bias even when no visual information is present. In other words, the acceptance of the ungrammatical locatives such as *Lucy filled water into the glass is not solely the artifact
of acceptability judgment task but also the product of content-oriented bias observed across both experimental studies.

Another important aspect of comparing the online and offline experimental outcomes is that each study offered a unique perspective on the developmental trends in L2 acquisition of locative semi-regularities. The eye-tracking findings demonstrated the relationship between L2 proficiency and the capacities of L2 learners to overcome the content-oriented bias in processing container (i.e., fill-kind) non-alternating structures: The more proficient the L2 participants were, the more native-like appeared their eye-movement patterns anticipating container arguments within the post-verbal temporal window of interest. On the other hand, the non-rejections of ungrammatical content-only (i.e., pour-kind) constructions on the acceptability judgment task (found by the absolute analysis) revealed similar associations between the L2 participants’ English proficiency and their likelihood of accepting the ill-formed constructions: The higher proficiency L2 learners were consistently less likely to judge as acceptable the ungrammatical content non-alternating locatives. Therefore, even though neither offline nor online investigation is immune to task-induced influences, jointly they offer a more complete and accurate outlook on the experimental findings.

The findings of Experiment 2 provide additional empirical evidence needed to reassess the validity of the alternative proposals concerning the acquisition of locative constructions in L2. Similar to the outcomes of Experiment 1, the results of Experiment 2 pose considerable challenges for the position advocated by Bley-Vroman and Joo (2001) and Joo (2003). It is highly unlikely that the target-language-like eye-movement behavior observed across all experimental groups reflects a general lack of representational means to acquire the target constructions or points to an “unprincipled” way by which L2 learners deal with the locative
semi-regularities. Rather, the observed patterns suggest that L2 learner are capable of learning the nuanced argument structure of locatives and do so in a highly consistent and systematic manner. Ceteris paribus, the eye-tracking outcomes provide support for Juffs’s (1996) parametric account. As predicted by this proposal, Chinese native speakers learning L2 English indeed demonstrated target-like performance in processing content-only structures and matched L1 group in their real-time perception of container-only constructions only at the most advanced English proficiency level. However, the similarities between L1 and L2 eye-movement outcomes cast doubts on the validity of the parametric claim. It is highly improbable that L1 English speakers’ content bias reflects the parametric settings of Mandarin Chinese or that Chinese native speakers learning L2 English remain unaffected by the content-oriented bias characteristic of the performance of the English native speaking group. Even if the L1 does affect the ways L2 English learners perceive locative constructions, such an influence cannot be the sole source of the experimental observations. A more general cognitive mechanism transcending the lexical parameter posited by Juffs (1996) is more likely to be responsible for the participants’ behavioral patterns observed in both experiments. Finally, the empirical findings of Experiment 2, similar to the outcomes of Experiment 1, challenge the initial formulation of the processing-based account: None of the experimental groups, even beginning L2 English learners displayed tendencies to treat locative non-alternants as alternating structures. On the other hand, the observed eye-movement patterns are sufficiently compatible with the revised version of the current proposal, but, at the same time, necessitate integration of the content-oriented bias into the processing-based account. The following section is especially dedicated to such integration.
Content-oriented bias in language processing and acquisition

The consistent pattern of experimental outcomes revealing asymmetry between content- and container-oriented constructions deserves a separate strand in the current discussion. All participants in both Experiment 1 and Experiment 2 demonstrated content-bias effects. In the acceptability judgment task, L2 participants as well as L1 English speakers were less likely to accept ill-formed content-only items (e.g., *poured the glass with water) in comparison to their more error-prone ratings of container non-alternating locatives (e.g., *filled water into the glass). Also, L2 learners displayed the statistically significant developmental growth of their preferences for well-formed over ill-formed realizations of content-only verbs (but not for container non-alternating test items). Moreover, the frequency effects demonstrated by the High L2 proficiency group were confined exclusively to their ratings of content non-alternating locatives. Even more pronounced were the influences of the content-bias in the eye-tracking outcomes of Experiment 2. In processing non-alternating constructions, this bias consistently facilitated predictions of the appropriate arguments for content-only locatives, and, correspondingly, hampered the anticipatory eye-fixations for container-only verbs.

Similar observations can be found across a broad spectrum of publications in the linguistic literature. For example, Kim et al. (1990) presented a survey examining how locative verbs are expressed structurally across different languages. Based on their analysis, the authors claimed that languages fall into two basic classes. One class including such languages as English, French, Spanish, Singapore Malay, and Classical Arabic subdivides locative verbs into content-/container-only and alternating subcategories. The second class of languages including Korean, Japanese, Chinese, Thai, and Turkish has a simpler argument realization pattern allowing all locative verbs to be expressed in content-oriented sentence frames (cf. Juffs (1996)), thus
subdividing locatives into two subcategories: content-only and alternating. None of the languages examined by the survey, however, was found to subdivide locatives into container-only and alternating categories even though such categorization would be a logically plausible option. A similar content-oriented bias was found in a neuroimaging study of language processing conducted by Christensen and Wallentin (2011). Most important for the current discussion are the fMRI outcomes which revealed increased activation of the left interior frontal gyrus (LIFG) during comprehension of container-oriented locative constructions (as compared to processing content-oriented stimuli), suggesting that such structures as *John filled the glass with water* were more difficult to process than their content-oriented counterparts (e.g., *John poured water into the glass*). The content-container asymmetry is also evident in L1 and L2 acquisition research. For example, Bowerman (1982) analyzed spontaneous speech recordings of her daughters’ developing language suggesting that most errors in locative constructions involved *fill*-type verbs (e.g., *filled water into the glass*) with only few instances in which *pour*-type locatives were overextended (e.g., *pour the glass with water*). The same conclusions come from experimental studies examining children learning L1 English. Gropen et al. (1991) tested young English learners’ semantic and syntactic knowledge of six locative verbs: *pour, dump, fill, empty, splash, and stuff*. The findings revealed the strong content-oriented bias on both comprehension and production measures. The young participants performed similar to adult English speakers on experimental tasks involving content non-alternating verbs (*to pour* and *to dump*), but, at the same time, were more error-prone in their interpretation and production of container-only (*to fill* and *to empty*) locatives (see also Kim et al.).
Finally, the same sort of outcomes is evident in L2 investigations. Both Juffs (1996) and Ambridge and Brandt (2013) found that fill-type non-alternants posed greater challenges for L2 learners in comparison to content non-alternating (pour-type) locative verbs.

How can we explain the content-oriented bias observed across such diverse language phenomena? Even though the differences between content- and container-oriented constructions have received considerable attention in the linguistic literature (e.g., see Rappaport Hovav & Levin, 2010; Tenny, 1994), the only approach which explicitly addressed the issue of the content-bias in language acquisition is Pinker’s semantic verb class hypothesis. Pinker (1989) argues that the primary source of the content-oriented bias is a greater perceptual salience of the ways substances move or are caused to move across space, and, correspondingly, a more abstract (less salient) nature of the resulting state changes. The differences in perceptual salience, therefore, explain why children more easily acquire the meanings of such verbs as to pour and to spill, and, as a result, make few errors in using these verbs in speech. At the same time, the increased sensitivity to the manners of motion also explains more error-prone usages of container-only locatives such as to fill and to cover: The problems arise from attributing unintended but perceptually salient senses to the container non-alternating locatives, such as assigning pouring-like manners to the mannerless verb to fill. For Pinker, however, such explanation is a way to spell out the causes of the developmental errors observed in children’s language learning. Once the proper verbal meanings are acquired, the argument structure errors cease to exist.

Even though applied linguists can apply the same logic to L2 development, a number of empirical findings preclude researchers and theorists from construing the content-oriented bias as merely a developmental stumbling block. In both experiments of the current project, L1 English
participants demonstrated the content-bias effects. The similar findings of Christensen and Wallentin’s (2011) brain imaging study suggest that the content-container asymmetry cannot be limited to the initial stages of language development but also operates in highly fluent language speakers.

My preferred solution of the content-oriented bias is fully compatible with Pinker’s (1989) proposals, but extents the language-cognition interaction beyond the processes of language acquisition. Essentially, I view the content bias as the effects of perceptual and motor representations associated with the conceptual structures of content and container non-alternating locative verbs. The specific manners by which substances move or are caused to move across space are the key elements of such verbs as to pour, to spill, and to drip. In learning locative verbs, language acquirers necessarily associate the perceptual (primarily visual) representations with the verbal conceptual structures. Also, given that these verbs denote specific actions, the motor representations (the motor structures underlying, for example, an act of pouring) become consistent mental coordinates of the verbal meanings. Therefore, the consistently reinstated sensory-motor structures of content non-alternating verbs become an integral part of their extended meanings. In contrast, container-only locatives, such as to fill, to cover, and to decorate can be associated with a range of different ways substances move or are caused to move, and the acts of filling, covering and decorating can be accomplished by different motor sequences. Moreover, even though there are some stable perceptual coordinates of container-only verbs, such as visually perceivable changes of locative receptacles, the prototypical motor associates of the verbs to fill, to cover, and to decorate might in fact be in conflict with the semantic focus of the container non-alternating locative. For instance, a prototypical action associated with filling a glass with water is first and foremost about exerting forces on water (or water containing
entities) rather than a glass, and in most cases involves letting it flow into a container in a pouring manner. Similarly the acts of covering or decorating start and proceed with an agent’s actions directed towards contents, which, in turn, change containers’ states.

The differences between sensory-motor associates of content- and container non-alternating locatives affect the ways we process and correspondingly acquire their argument structures. For the pour-type verbs, the distinctive perceptual coordinates and the fully aligned motor representations facilitate the caused motion interpretations of these verbs serving as primes in language production or comprehension. On the other hand, a broad range of perceptual structures associated with the fill-type verbs and prototypical motor representations incongruent with container state changes in effect interfere with construction of container-oriented conceptual representations by enhancing competition between containers and contents during processing of container non-alternating verbs.

Therefore, my account of the content-oriented bias retains Pinker’s (1989) claim about the effects of perceptually salient aspects of locative events in language learning. At the same time, I ascribe a greater role to perception and action in language processing and view the content-bias in language acquisition as a revealing instantiation of how sensory-motor systems interact with language in general.

The revised processing-based account

An apparent conclusion following from the review of the experimental outcomes is that any property theory relying on specifications of what L2 learners know (or fail to know) at different stages in L2 development would face considerable challenges in explaining the observed behavioral patterns. In contrast, the empirical outcomes can be naturally understood
within a more dynamic framework allowing an integration of the property-theoretical constructs with the mechanisms explaining changes in the linguistic system.

The central thesis of the processing-based account advocated by the current project is that the development of the locative constructions in L2 is primarily the matter of moment-by-moment mental processes underlying our capacities to understand and produce language. It treats overgeneralizations and retreat from the overly general argument realization patterns as a natural phenomenon arising from language learners’ attempts to re-represent (process) complex and fast-paced environmental stimuli. Couched within the MOGUL framework, the current proposal is best understood as an instantiation of the acquisition-by-processing hypothesis in an analysis of a particular grammatical construction in L2 learning. In most general terms, the account views linguistic development as a byproduct of information processing events (language learners’ linguistic experiences) in which a set of independent but interrelated cognitive expert systems (i.e., modules) construct mental structures continuously striving to find the best fit (and/or best coordinates) for the representations built by the adjacent modules. Whenever we hear/read a single word or a complex linguistic structure, our experiences are determined and limited by the nature and capacities of the dedicated expert systems (phonological, syntactic, conceptual/semantic, visual, affective, etc.) to incrementally construct and coordinate mental representations ultimately creating chains of mental structures constituting our knowledge of a language. The chains of mental structures constructed for the purpose of comprehension become strengthened each time they are reused in processing gradually gaining greater mnemonic accessibility. Such representational chains, rather than environmental entities (i.e., linguistic evidence) from which linguistic generalizations can be extracted, therefore, are the primary
Theoretical constructs the current perspective relies on in the discussion of the locative phenomena.

The locative semi-regularities stem from, on the one hand, our flexibility in conceptualizing locative events, and, on the other hand, particular ways a language happens to express specific aspects of these complex, dynamic interrelationships. What this means in the developmental terms is that the availability of alternative interpretations, and correspondingly, alternative ways of encoding these interpretations structurally is a default option for all locative verbs. For a subset of locative verbs such default is fully realized: The regularly alternating locative verbs such to load, to spray, and to spread, are a part of L2 language learners’ linguistic experiences. By hearing sentences containing these verbs (i.e., building PS⇔SS⇔CS⇔…) representational chains) across multiple communicative contexts, L2 language acquirers develop the means to flexibly express both interpretations of locative events.

More important for the current discussion of the semi-regularities is how L2 learners deal with the irregular locatives disallowing the default flexibility. In a nutshell, the current account views constraining potential overgeneralizations as strengthening of attested argument realizations patterns (representational chains) by recurrently (re)constructing them in processing. This necessarily involves learning the precise semantics of the lexical forms as well as statistically driven mnemonic reinforcement (increase of mnemonic accessibility) of the form-meaning correspondences experienced by language learners.

The alternative (unattested) locative construals/structural realizations, however, are never expelled from language learners’ grammar but gradually become less accessible as the attested constructions growth in their availability. The divergences in accessibility of attested versus unattested argument realizations become apparent at the early stages of L2 development. As soon
as L2 learners become familiar with non-alternating locative verbs, they recognize that pouring coffee... and filling a cup... sound better than *pouring a cup and *filling coffee. Such divergent trends do not necessarily reach their extremes at any point in the development. What this means is that even though lower proficiency L2 learners are naturally more likely to overextend the non-alternating locatives in comparison to advanced L2 or L1 English speaker, the highly fluent English users might still demonstrate such overgeneralizations in either production or comprehension. The overextensions in production are likely to be relatively rare and observed in either highly constrained contexts priming L2 speakers to produce the ill-formed sentences or appear as intentionally creative uses of the verbs (in Pinkers’ term “one-shot innovations” such as when adult L1 English speakers produce He squeezed the fish with lemon juice p. 374). In comprehension, especially in artificial contexts of experimental studies, the errors are expected to be more frequent (in comparison to production) but are likely to decrease gradually with higher proficiency in the L2. Finally, the developmental trends and the patterns of overextensions are likely to vary within two categories of non-alternating locatives with frequently reinstated/reprocessed locative verbs showing steeper development trends and fewer overextensions in comparison with their relatively less frequent counterparts.

**Concluding remarks**

From its inception, the current project sought to clarify the issues surrounding learning locative constructions in L2. It started from a recognition of inconsistencies among the theoretical ideas and empirical observations advanced by a small number of L2 publications. Despite their pioneering roles, these investigations offered rather isolated accounts of locative phenomena providing no common grounds for any coherent views on either processes or outcomes of learning locative semi-regularities in L2. The current processing-based proposal is
the product of theoretical and empirical reexamination of the conflicting accounts. It offers a novel theoretical view accommodating complex behavioral patterns in acceptability judgments and online-processing of locative constructions. At the same time, it is based on limited experimental evidence and, therefore, requires further empirical data to substantiate or refute its claims and underlying assumptions.

First of all, in order to provide a better description of learning the target constructions, we need spontaneous production data showing how locative verbs emerge and are used in either speech or writing across the earliest stages of L2 development. Similarly, it is imperative to obtain more accurate estimates of what language L2 learners are exposed to, and correspondingly, how the externally available linguistic samples affect or fail to affect L2 learners’ developing system. Thus, longitudinal observations, corpus-based studies, as well as experimentations focusing on beginning L2 learners are the necessary means to eliminate the current blind spots in the empirical evidence.

Also, the processing-based account hypothesizes involvement of semantic and statistical learning mechanisms in the development of the locative semi-regularities relying, first and foremost, on theoretical argumentation. Even though it is quite challenging (or probably impossible) to disentangle the specific effects of two learning mechanisms, the current proposal can be used to design experimental (or quasi-experimental) investigations examining how cross-situational learning (i.e., contextual and sensory-motor factors) and frequencies of exposure affect L2 learners’ perception of locatives. These studies can explore novel or familiar locative verbs and can be done with L2 learners of different proficiencies providing evidence for both theoretical and educational purposes.
The main focus on the current proposal is on the psycholinguistic dimension of L2 development. Therefore, a number of postulated processes and outcomes, such as context and task-based overextensions, dependence of overgeneralizations on mnemonic availability of unintended interpretations, content-bias as the effect of sensory-motor coordinates, etc., require focused experimental evaluation. Such methodologies as self-paced reading, eye-tracking, and structural priming/interference are the primary tools which can be used to obtain empirical evidence necessary for refining or refuting the current processing-based proposal. The use of psycholinguistic experimentation is also necessary to clarify the extent to which the native language of L2 learners or bi/multilingualism affects perception and learning novel linguistic constructions distinguishing them from non-linguistic influences in language development and use.
Appendix A: Lexical characteristics of the target locative verbs by categories

Table 12. *Means and standard deviations in parenthesis of locative verbs’ characteristics*

<table>
<thead>
<tr>
<th></th>
<th>Non-alternating</th>
<th>Alternating</th>
<th>$F$ (3, 20)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Content Only</td>
<td>Container Only</td>
<td>Content</td>
<td>Container</td>
</tr>
<tr>
<td>Freq $\text{COCA}$</td>
<td>4.06 (0.40)</td>
<td>3.90 (0.84)</td>
<td>3.80 (0.40)</td>
<td>3.82 (0.33)</td>
</tr>
<tr>
<td>PLD20</td>
<td>1.18 (0.21)</td>
<td>1.51 (0.66)</td>
<td>1.72 (0.55)</td>
<td>1.18 (0.19)</td>
</tr>
<tr>
<td>N phone</td>
<td>3.50 (0.55)</td>
<td>4.00 (1.63)</td>
<td>4.67 (1.91)</td>
<td>3.50 (1.43)</td>
</tr>
<tr>
<td>RT $\text{LEX DEC}$</td>
<td>598.11 (28.40)</td>
<td>655.42 (63.61)</td>
<td>648.64 (45.79)</td>
<td>610.91 (44.93)</td>
</tr>
<tr>
<td>Acc $\text{LEX DEC}$</td>
<td>0.98 (0.02)</td>
<td>0.97 (0.03)</td>
<td>0.97 (0.03)</td>
<td>0.97 (0.04)</td>
</tr>
</tbody>
</table>

*Notes.* Freq $\text{COCA}$ = logarithmically transformed frequency extracted from the COCA corpus (Davis, 2008); PLD20 = phonological Levenshtein distance 20 (measurement of phonological neighborhood); N phone = number of phonemes, RT $\text{LEX DEC}$ = mean reaction time of lexical decision in msec; Acc $\text{LEX DEC}$ = mean accuracy of lexical decision. All lexical characteristics come from the English Lexicon Project database (Balota et al, 2007).
Appendix B: Sentence stimuli for Experiment 1

A. Content alternating sentence stimuli

<table>
<thead>
<tr>
<th>Content-oriented sentence frame</th>
<th>Container-oriented sentence frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>James spread butter onto the toast.</td>
<td>James spread the toast with butter.</td>
</tr>
<tr>
<td>Lisa sprinkled cheese on the pizza.</td>
<td>Lisa sprinkled the pizza with cheese.</td>
</tr>
<tr>
<td>Chris smeared jam on the cookie.</td>
<td>Chris smeared the cookie with jam.</td>
</tr>
<tr>
<td>Jennifer splashed milk onto the cereal.</td>
<td>Jennifer splashed the cereal with milk.</td>
</tr>
<tr>
<td>The baker sprayed oil on the pan.</td>
<td>The baker sprayed the pan with oil.</td>
</tr>
<tr>
<td>The cleaner piled dishes onto the counter.</td>
<td>The cleaner piled the counter with dishes.</td>
</tr>
</tbody>
</table>

B. Container alternating sentence stimuli

<table>
<thead>
<tr>
<th>Content-oriented sentence frame</th>
<th>Container-oriented sentence frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ron loaded spaghetti on the plate.</td>
<td>Ron loaded the plate with spaghetti.</td>
</tr>
<tr>
<td>Michelle packed meat into the freezer.</td>
<td>Michelle packed the freezer with meat.</td>
</tr>
<tr>
<td>The grandmother stuffed mushrooms into the turkey.</td>
<td>The grandmother stuffed the turkey with mushrooms.</td>
</tr>
<tr>
<td>The postal carrier jammed the mailbox with letters.</td>
<td>The postal carrier jammed letters into the mailbox.</td>
</tr>
<tr>
<td>The shopkeeper stocked cans on the shelf.</td>
<td>The shopkeeper stocked the shelf with cans.</td>
</tr>
<tr>
<td>The tourist crammed clothes into the suitcase.</td>
<td>The tourist crammed the suitcase with clothes.</td>
</tr>
</tbody>
</table>

C.1 Content non-alternating sentence stimuli (Argument set A)

<table>
<thead>
<tr>
<th>Content-oriented sentence frame</th>
<th>Container-oriented sentence frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richard dripped juice onto the fish.</td>
<td>*Richard dripped the fish with juice.</td>
</tr>
<tr>
<td>Maria dumped sauce onto the pasta.</td>
<td>*Maria dumped the pasta with sauce.</td>
</tr>
<tr>
<td>The photographer pinned pictures onto the wall.</td>
<td>*The photographer pinned the wall with pictures.</td>
</tr>
<tr>
<td>The bartender poured water into the glass.</td>
<td>*The bartender poured the glass with water.</td>
</tr>
<tr>
<td>The reporter spilled coffee on the t-shirt.</td>
<td>*The reporter spilled the t-shirt with coffee.</td>
</tr>
<tr>
<td>Ken shook salt into the soup.</td>
<td>*Ken shook the soup with salt.</td>
</tr>
</tbody>
</table>
### C.2 Content non-alternating sentence stimuli (Argument set B)

<table>
<thead>
<tr>
<th>Content-oriented sentence frame</th>
<th>Container-oriented sentence frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richard dripped honey onto the pancake.</td>
<td>*Richard dripped the pancake with honey.</td>
</tr>
<tr>
<td>Maria dumped gravy onto the steak.</td>
<td>*Maria dumped the steak with gravy.</td>
</tr>
<tr>
<td>The photographer pinned a flower onto the jacket.</td>
<td>*The photographer pinned the jacket with a flower.</td>
</tr>
<tr>
<td>The bartender poured tea into the cup.</td>
<td>*The bartender poured the cup with tea.</td>
</tr>
<tr>
<td>The reporter spilled wine on the jeans.</td>
<td>*The reporter spilled the jeans with wine.</td>
</tr>
<tr>
<td>Ken shook pepper into the stew.</td>
<td>*Ken shook the stew with pepper.</td>
</tr>
</tbody>
</table>

### D.1 Container non-alternating sentence stimuli (Argument set A)

<table>
<thead>
<tr>
<th>Content-oriented sentence frame</th>
<th>Container-oriented sentence frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Laura coated gravy on the steak.</td>
<td>Laura coated the steak with gravy.</td>
</tr>
<tr>
<td>*The maid filled tea into the cup.</td>
<td>The maid filled the cup with tea.</td>
</tr>
<tr>
<td>*The writer covered honey on the pancake.</td>
<td>The writer covered the pancake with honey.</td>
</tr>
<tr>
<td>*Susan stained wine on the jeans.</td>
<td>Susan stained the jeans with wine.</td>
</tr>
<tr>
<td>*Charles seasoned pepper onto the stew.</td>
<td>Charles seasoned the stew with pepper.</td>
</tr>
<tr>
<td>*The flight attendant decorated a flower onto the jacket.</td>
<td>The flight attendant decorated the jacket with a flower.</td>
</tr>
</tbody>
</table>

### D.2 Container non-alternating sentence stimuli (Argument set B)

<table>
<thead>
<tr>
<th>Content-oriented sentence frame</th>
<th>Container-oriented sentence frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>*The writer covered juice on the fish.</td>
<td>The writer covered the fish with juice.</td>
</tr>
<tr>
<td>*Laura coated sauce on the pasta.</td>
<td>Laura coated the pasta with sauce.</td>
</tr>
<tr>
<td>*The maid filled water into the glass.</td>
<td>The maid filled the glass with water.</td>
</tr>
<tr>
<td>*Susan stained coffee on the t-shirt.</td>
<td>Susan stained the t-shirt with coffee.</td>
</tr>
<tr>
<td>*Charles seasoned salt into the soup.</td>
<td>Charles seasoned the soup with salt.</td>
</tr>
<tr>
<td>*The flight attendant decorated pictures onto the wall.</td>
<td>The flight attendant decorated the wall with pictures.</td>
</tr>
</tbody>
</table>
F. Practice sentences

Mary had a little lamb.

The car the dog the cat chased.

There is a great restaurant in the city.

The notebook the candle burnt.

Sam recites poems as well as playing the piano.

In China, wine is served in small cups, and in Turkey, coffee.

Someone apparently vanished my wallet.

After the exams are the time to relax.
Appendix C: A sample of the acceptability judgment trial

The maid filled water into the glass.
Appendix D: L1 English speaker background questionnaire

What is your gender?

☐ Male ☐ Female

How old are you?

Please check your highest education level

☐ I am an undergraduate student
☐ Undergraduate Degree
☐ I am a Master's student
☐ Master’s Degree
☐ I am a PhD student
☐ PhD Degree

Do you speak any additional language?

☐ Yes ☐ No

Please indicate the additional language you know


Appendix E: L2 learner background questionnaire

What is your gender?

☐ Male  ☐ Female

How old are you?


At what age did you start learning English?


At what age did you come to the USA?


Please check your **highest** education level

- ☐ I am an undergraduate student
- ☐ Undergraduate Degree
- ☐ I am a Master’s student
- ☐ Master’s Degree
- ☐ I am a PhD student
- ☐ PhD Degree

What percentage of the time do you spend communicating (speaking, writing, listening and speaking) in English and Chinese?  *(Note: Total must sum to 100)*

<table>
<thead>
<tr>
<th>Language</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td></td>
</tr>
<tr>
<td>English</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>0</td>
</tr>
</tbody>
</table>
Appendix F: Vocabulary Knowledge Scale Test

Directions: Please consider each of the words below and check the box that best describes how well you know and understand the word

**to drip**

☐ I don't remember having seen this word before.

☐ I have seen this word before, but I don't know what it means.

☐ I have seen this word before, and I know what it means.

☐ I know this word well and can use it in a sentence.
Appendix G: Timing information for acceptability judgment task

Table 13. *Means and standard deviations for timings of participants’ responses in Experiment 1*

<table>
<thead>
<tr>
<th>Group</th>
<th>SF</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Non Alternating</td>
<td>Alternating</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Content</td>
<td>Container</td>
<td>Content</td>
<td>Container</td>
</tr>
<tr>
<td>L1</td>
<td>Container</td>
<td>5.34 (2.22)</td>
<td>4.54 (3.36)</td>
<td>4.67 (2.21)</td>
<td>4.65 (3.68)</td>
</tr>
<tr>
<td></td>
<td>Content</td>
<td>5.65 (3.83)</td>
<td>6.81 (4.26)</td>
<td>5.28 (3.30)</td>
<td>5.50 (3.58)</td>
</tr>
<tr>
<td>High L2</td>
<td>Container</td>
<td>7.60 (3.23)</td>
<td>6.49 (3.10)</td>
<td>7.23 (3.18)</td>
<td>7.06 (3.30)</td>
</tr>
<tr>
<td></td>
<td>Content</td>
<td>8.31 (4.70)</td>
<td>9.24 (4.25)</td>
<td>7.43 (2.98)</td>
<td>8.97 (4.77)</td>
</tr>
<tr>
<td>Intermediate L2</td>
<td>Container</td>
<td>7.60 (4.57)</td>
<td>6.51 (4.37)</td>
<td>6.66 (3.64)</td>
<td>6.68 (3.59)</td>
</tr>
<tr>
<td></td>
<td>Content</td>
<td>8.39 (6.41)</td>
<td>10.21 (7.03)</td>
<td>8.57 (5.54)</td>
<td>9.27 (5.77)</td>
</tr>
<tr>
<td>Low L2</td>
<td>Container</td>
<td>6.62 (3.23)</td>
<td>6.54 (4.09)</td>
<td>7.55 (7.67)</td>
<td>6.34 (3.25)</td>
</tr>
<tr>
<td></td>
<td>Content</td>
<td>8.27 (4.01)</td>
<td>9.32 (5.06)</td>
<td>7.73 (4.03)</td>
<td>8.68 (5.82)</td>
</tr>
</tbody>
</table>

*Note:* SF= sentence frame. Means and standard deviations (in parenthesis) are in seconds.
Appendix H: Summary of data reduction for Experiment 1

Table 14. Percentage of discarded test items in Experiment 1 arranged by experimental groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Fast</th>
<th>Unknown</th>
<th>Outliers</th>
<th>Other</th>
<th>Removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low L2</td>
<td>1.68</td>
<td>22.45</td>
<td>1.22</td>
<td>0.00</td>
<td>25.35</td>
</tr>
<tr>
<td>Intermediate L2</td>
<td>0.42</td>
<td>24.17</td>
<td>1.25</td>
<td>0.00</td>
<td>25.83</td>
</tr>
<tr>
<td>High L2</td>
<td>0.13</td>
<td>16.94</td>
<td>1.68</td>
<td>3.23</td>
<td>21.98</td>
</tr>
<tr>
<td>L1 English</td>
<td>0.00</td>
<td>0.00</td>
<td>2.13</td>
<td>0.00</td>
<td>2.13</td>
</tr>
</tbody>
</table>

*Note:* Fast = responses submitted prior to 1500 ms cutoff point; Unknown = test items containing verbs reported as unknown by L2 learners in the VKS survey; Outliers = acceptability ratings below or above 2.5 standard deviations of the respective group means; Other = data loss due to a recording failure.
Appendix I: Trials removed after dichotomization of participants’ responses in Experiment 1

Table 15. Percentage of “uncertain” responses in Experiment 1 organized by proficiency groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Content Only</th>
<th>Container Only</th>
<th>Content Alt</th>
<th>Container Alt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Content</td>
<td>Container</td>
<td>Content</td>
<td>Container</td>
</tr>
<tr>
<td>L1</td>
<td>2.32</td>
<td>10.90</td>
<td>10.90</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1.32</td>
<td>3.36</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>High</td>
<td>3.68</td>
<td>9.09</td>
<td>6.71</td>
<td>3.23</td>
</tr>
<tr>
<td></td>
<td>5.34</td>
<td>6.02</td>
<td>4.76</td>
<td>5.19</td>
</tr>
<tr>
<td>Intermediate</td>
<td>7.74</td>
<td>11.32</td>
<td>10.73</td>
<td>2.94</td>
</tr>
<tr>
<td></td>
<td>7.75</td>
<td>4.55</td>
<td>4.94</td>
<td>4.73</td>
</tr>
<tr>
<td>Low</td>
<td>8.05</td>
<td>7.30</td>
<td>13.11</td>
<td>4.52</td>
</tr>
<tr>
<td></td>
<td>8.40</td>
<td>7.91</td>
<td>6.59</td>
<td>11.18</td>
</tr>
</tbody>
</table>

Notes. The values indicate percentage of trials on which participants chose option “4”. These values were computed separately for each type of locative verbs used in content- and container-oriented sentence frames. The column Total represents a total percentage of “uncertain” trial for each experimental group.

Additional data analysis

The participants were uncertain about acceptability of the test items on 5.66% of the trials. An additional data analysis examined how likely participants (of different proficiency levels) were to become uncertain when they saw the target verbs (of different categories) in the two alternative locative constructions. To run the analysis, I recoded the outcome variable by collapsing the “accept” and “reject” responses into a new category “certain” and defining responses “4” as “uncertain”. I used logistic regression modeling, described in the Statistical Analysis section of Experiment 1. The results showed that participants’ degree of “uncertainty” was significantly higher for the ungrammatical locative constructions but did not differ between two argument realization patterns of the alternating verbs. Specifically, ungrammatical content-only locatives across all experimental groups (L1 English, $\beta = 2.71$, $SE = 0.54$, $p < .00001$; High
L2, $\beta = 0.92, SE = 0.48, p = .058$; Intermediate L2, $\beta = 1.38, SE = 0.45, p = .002$; and Low L2, $\beta = 1.23, SE = 0.39, t = 3.14, p = .0016$), as well as ungrammatical container-only constructions for two higher proficiency level groups (L1 English, $\beta = 1.39, SE = 0.35, p = .00001$; High L2, $\beta = 0.82, SE = 0.44, p = .065$) were more likely to evoke the “uncertain” responses, in comparison to their grammatical counterparts. Compared to native speakers, all L2 groups were more likely to demonstrate the “uncertainty” for the grammatical but not ungrammatical structures. Therefore, the uncertainty in the participants’ responses were primarily driven by the ungrammatical locative structures and further modulated by the English proficiency level.
Appendix J: Summary of logistic regression analyses

Table 16. *Summary of logistic regression model for the content non-alternating locative stimuli*

<table>
<thead>
<tr>
<th></th>
<th>Estimate ($\beta$)</th>
<th>Odds Ratio</th>
<th>Estimate</th>
<th>$SE$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>L1 GROUP; CONTAINER-ORIENTED</td>
<td>0.38</td>
<td>-0.97***</td>
<td>0.13</td>
</tr>
<tr>
<td>Sentence Frame</td>
<td>CONTENT -ORIENTED</td>
<td>605.16</td>
<td>6.41***</td>
<td>0.80</td>
</tr>
<tr>
<td>Proficiency HIGH</td>
<td></td>
<td>2.16</td>
<td>0.77***</td>
<td>0.22</td>
</tr>
<tr>
<td>Proficiency INTERMEDIATE</td>
<td></td>
<td>2.96</td>
<td>1.087***</td>
<td>0.22</td>
</tr>
<tr>
<td>Proficiency LOW</td>
<td></td>
<td>4.14</td>
<td>1.42***</td>
<td>0.21</td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency HIGH</td>
<td></td>
<td>0.07</td>
<td>-2.64**</td>
<td>0.94</td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency INTERMEDIATE</td>
<td></td>
<td>0.01</td>
<td>-4.49***</td>
<td>0.85</td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency LOW</td>
<td></td>
<td>0.01</td>
<td>-5.15***</td>
<td>0.84</td>
</tr>
</tbody>
</table>

*Notes.* ** $p < 0.1$; *** $p < .001$; SF = Sentence Frame; $SE$ = Standard Error

Table 17. *Summary of logistic regression model for container non-alternating locative stimuli*

<table>
<thead>
<tr>
<th></th>
<th>Estimate ($\beta$)</th>
<th>Odds Ratio</th>
<th>Estimate</th>
<th>$SE$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>L1 GROUP; CONTENT-ORIENTED</td>
<td>0.94</td>
<td>-0.06</td>
<td>0.12</td>
</tr>
<tr>
<td>Sentence Frame</td>
<td>CONTAINER-ORIENTED</td>
<td>225.94</td>
<td>5.42***</td>
<td>0.75</td>
</tr>
<tr>
<td>Proficiency HIGH</td>
<td></td>
<td>1.71</td>
<td>0.54**</td>
<td>0.20</td>
</tr>
<tr>
<td>Proficiency INTERMEDIATE</td>
<td></td>
<td>1.75</td>
<td>0.56**</td>
<td>0.20</td>
</tr>
<tr>
<td>Proficiency LOW</td>
<td></td>
<td>1.92</td>
<td>0.65**</td>
<td>0.20</td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency HIGH</td>
<td></td>
<td>0.06</td>
<td>-2.82***</td>
<td>0.84</td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency INTERMEDIATE</td>
<td></td>
<td>0.03</td>
<td>-3.53***</td>
<td>0.81</td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency LOW</td>
<td></td>
<td>0.04</td>
<td>-3.14***</td>
<td>0.83</td>
</tr>
</tbody>
</table>

*Notes.* ** $p < 0.1$; *** $p < .001$; SF = Sentence Frame; $SE$ = Standard Error
### Table 18. *Summary of logistic regression model for the content alternating locative stimuli*

<table>
<thead>
<tr>
<th></th>
<th>Estimate (β)</th>
<th>Odds Ratio</th>
<th>Estimate</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept L1 GROUP; CONTAINER-ORIENTED</td>
<td>14.29</td>
<td>2.66***</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Sentence Frame CONTENT-ORIENTED</td>
<td>15.58</td>
<td>2.75***</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>Proficiency HIGH</td>
<td>0.12</td>
<td>-2.16***</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>Proficiency INTERMEDIATE</td>
<td>0.14</td>
<td>-1.99***</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Proficiency LOW</td>
<td>0.17</td>
<td>-1.78***</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency HIGH</td>
<td>0.34</td>
<td>-1.08</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency INTERMEDIATE</td>
<td>0.20</td>
<td>-1.60*</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency LOW</td>
<td>0.22</td>
<td>-1.49</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>

*Notes. *p < .05; ** p < 0.1; *** p < .001; SF = Sentence Frame, SE = Standard Error*

### Table 19. *Summary of logistic regression model for the container alternating locative stimuli*

<table>
<thead>
<tr>
<th></th>
<th>Estimate (β)</th>
<th>Odds Ratio</th>
<th>Estimate</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept L1 GROUP; CONTAINER-ORIENTED</td>
<td>706.55</td>
<td>6.56***</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Sentence Frame CONTENT-ORIENTED</td>
<td>0.75</td>
<td>-0.29</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Proficiency HIGH</td>
<td>0.02</td>
<td>-4.09***</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>Proficiency INTERMEDIATE</td>
<td>0.01</td>
<td>-4.33***</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>Proficiency LOW</td>
<td>0.01</td>
<td>-4.69***</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency HIGH</td>
<td>0.39</td>
<td>-0.94</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency INTERMEDIATE</td>
<td>0.42</td>
<td>-0.87</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>SF CONTENT-ORIENTED * Proficiency LOW</td>
<td>0.47</td>
<td>-0.75</td>
<td>0.89</td>
<td></td>
</tr>
</tbody>
</table>

*Notes. *p < .05; ** p < 0.1; *** p < .001; SF = Sentence Frame, SE = Standard Error*
Appendix K: Summary of mixed-effect model for L2 participants difference scores

Table 20. Mixed-effect model summary for non-alternating locatives’ preference scores

<table>
<thead>
<tr>
<th></th>
<th>Estimate (β)</th>
<th>SE</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept LOW L2; CONTENT-ORIENTED</td>
<td>1.09**</td>
<td>0.34</td>
<td>3.18</td>
</tr>
<tr>
<td>Proficiency INTERMEDIATE</td>
<td>0.57</td>
<td>0.38</td>
<td>1.49</td>
</tr>
<tr>
<td>Proficiency HIGH</td>
<td>1.12**</td>
<td>0.39</td>
<td>2.84</td>
</tr>
<tr>
<td>Verb Type CONTAINER</td>
<td>0.29</td>
<td>0.47</td>
<td>0.63</td>
</tr>
<tr>
<td>Frequency (COCA)</td>
<td>-0.39</td>
<td>0.73</td>
<td>-0.54</td>
</tr>
<tr>
<td>Proficiency INTERMEDIATE * Verb Type CONTAINER</td>
<td>-0.54</td>
<td>0.50</td>
<td>-1.08</td>
</tr>
<tr>
<td>Proficiency HIGH * Verb Type CONTAINER</td>
<td>-0.59</td>
<td>0.52</td>
<td>-1.18</td>
</tr>
<tr>
<td>Proficiency INTERMEDIATE * Frequency (COCA)</td>
<td>0.89</td>
<td>0.64</td>
<td>1.39</td>
</tr>
<tr>
<td>Proficiency HIGH * Frequency (COCA)</td>
<td>2.07**</td>
<td>0.66</td>
<td>3.12</td>
</tr>
<tr>
<td>Verb Type CONTAINER * Frequency (COCA)</td>
<td>0.52</td>
<td>0.81</td>
<td>0.64</td>
</tr>
<tr>
<td>Proficiency INTERMEDIATE * Verb Type CONTAINER * Frequency (COCA)</td>
<td>-0.78</td>
<td>0.71</td>
<td>-1.10</td>
</tr>
<tr>
<td>Proficiency HIGH * Verb Type CONTAINER * Frequency (COCA)</td>
<td>-2.12**</td>
<td>0.73</td>
<td>-2.90</td>
</tr>
</tbody>
</table>

Notes. ** p < 0.1
Appendix L: Sentence stimuli for Experiment 2

Content alternating sentence stimuli

<table>
<thead>
<tr>
<th>Content-oriented sentence frame</th>
<th>Container-oriented sentence frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>James spread the butter onto the toast.</td>
<td>James spread the toast with butter.</td>
</tr>
<tr>
<td>Lisa sprinkled the cheese on the pizza.</td>
<td>Lisa sprinkled the pizza with cheese.</td>
</tr>
<tr>
<td>Chris smeared the jam on the cookie.</td>
<td>Chris smeared the cookie with jam.</td>
</tr>
<tr>
<td>Jennifer splashed the milk onto the cereal.</td>
<td>Jennifer splashed the cereal with milk.</td>
</tr>
<tr>
<td>The baker sprayed the oil on the pan.</td>
<td>The baker sprayed the pan with oil.</td>
</tr>
<tr>
<td>The cleaner piled the dishes onto the counter.</td>
<td>The cleaner piled the counter with dishes.</td>
</tr>
</tbody>
</table>

Container alternating sentence stimuli

<table>
<thead>
<tr>
<th>Content-oriented sentence frame</th>
<th>Container-oriented sentence frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ron loaded the spaghetti on the plate.</td>
<td>Ron loaded the plate with spaghetti.</td>
</tr>
<tr>
<td>Michelle packed the meat into the freezer.</td>
<td>Michelle packed the freezer with meat.</td>
</tr>
<tr>
<td>The grandmother stuffed the mushrooms into the turkey.</td>
<td>The grandmother stuffed the turkey with mushrooms.</td>
</tr>
<tr>
<td>The postal carrier jammed the letters into the mailbox.</td>
<td>The postal carrier jammed the mailbox with letters.</td>
</tr>
<tr>
<td>The shopkeeper stocked the cans on the shelf.</td>
<td>The shopkeeper stocked the shelf with cans.</td>
</tr>
<tr>
<td>The tourist crammed the clothes into the suitcase.</td>
<td>The tourist crammed the suitcase with clothes.</td>
</tr>
</tbody>
</table>
### Non-alternating target sentences

<table>
<thead>
<tr>
<th>Container non-alternating (Argument set A)</th>
<th>Content non-alternating (Argument set A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The cook covered the fish with juice.</td>
<td>The cook dripped the juice onto the fish.</td>
</tr>
<tr>
<td>The cook stained the t-shirt with coffee.</td>
<td>The cook spilled the coffee on the t-shirt.</td>
</tr>
<tr>
<td>The manager filled the glass with water.</td>
<td>The manager poured the water into the glass.</td>
</tr>
<tr>
<td>The manager seasoned the soup with salt.</td>
<td>The manager shook the salt into the soup.</td>
</tr>
<tr>
<td>The waiter coated the pasta with sauce.</td>
<td>The waiter dumped the sauce onto the pasta.</td>
</tr>
<tr>
<td>The waiter decorated the wall with pictures.</td>
<td>The waiter pinned the pictures onto the wall.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Container non-alternating (Argument set B)</th>
<th>Content non-alternating (Argument set B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The cook covered the pancake with honey.</td>
<td>The cook dripped the honey onto the pancake.</td>
</tr>
<tr>
<td>The cook stained the jeans with wine.</td>
<td>The cook spilled the wine on the jeans.</td>
</tr>
<tr>
<td>The manager filled the cup with tea.</td>
<td>The manager poured the tea into the cup.</td>
</tr>
<tr>
<td>The manager seasoned the stew with pepper.</td>
<td>The manager shook the pepper into the stew.</td>
</tr>
<tr>
<td>The waiter coated the steak with gravy.</td>
<td>The waiter dumped the gravy onto the steak.</td>
</tr>
<tr>
<td>The waiter decorated the jacket with the flower.</td>
<td>The waiter pinned the flower onto the jacket.</td>
</tr>
</tbody>
</table>

### Distractor sentences

| The cook shattered the mirror with the bat. | The cook pitched the pebble into the pond. |
| The cook broke the window with the stone.  | The cook flipped the coin into the fountain. |
| The cook scraped the pot with the spoon.   | The cook threw the dice onto the barrel. |
| The cook split the log with the axe.      | The cook baked the cake in the oven. |
| The manager smashed the watch with the hammer. | The manager roasted the chestnuts in the fire. |
| The manager swept the floor with the broom. | The manager boiled the sweet corn on the stove. |
| The manager crushed the ice with the rolling pin. | The manager fried the onions in the skillet. |
| The manager scrubbed the tub with the brush. | The waiter tossed the newspapers in the bin. |
| The manager wiped the glasses with the cloth. | The waiter barbecued the pork chops on the grill. |
| The waiter scratched the car with the key.  | The waiter steamed the broccoli in the microwave. |
| The waiter polished the boots with the rag. | The waiter flung the phone on the couch. |
| The waiter cracked the egg with the knife.  | The waiter slapped the cards on the table. |
### Attention grabbers

**Auditory Stimuli**

| The cook dragged **the boat** on the shore. | desk chair |
| The cook dried the **sweater** in the sun. | bed sheet |
| The cook kicked the **fence** with the foot. | dog |
| The cook pricked the **balloon** with the needle. | pillow |
| The cook rode the **horse** in the field. | cow |
| The cook trimmed the **beard** with the scissors. | mustache |
| The cook walked the **dog** in the park. | cat |
| The cook wrote the **note** with the pencil. | book |
| The manager carried the **vase** in the bedroom. | lamp |
| The manager cut the **finger** with the saw. | knee |
| The manager drew the **tree** with the crayons. | sculpture |
| The manager drove the **bus** on the highway. | tricycle |
| The manager fed the **cow** with the hay. | horse |
| The manager found the **postcard** in the drawer. | wallet |
| The manager grew the **pumpkin** in the garden. | cucumber |
| The manager squeezed the **orange** with the juicer. | apple |
| The waiter caught the **butterfly** with the net. | rabbit |
| The waiter dropped the **lamp** on the carpet. | vase |
| The waiter fixed the **TV** with the screwdriver. | radio |
| The waiter flew the **kite** in the yard. | airplane |
| The waiter poked the **elephant** with the stick. | hippopotamus |
| The waiter pushed the **desk** in the corner. | chair |
| The waiter washed the **plate** with the soap. | hands |
| The waiter watched the **stars** in the sky. | moon |
Appendix M: Sample images used in Experiment 2
Appendix N: Summary of the growth curve regression analyses by proficiency level

Table 21. Summary of the grow curve regression results for L1 English group

<table>
<thead>
<tr>
<th></th>
<th>Estimate ($\beta$)</th>
<th>SE</th>
<th>$t$</th>
<th>$\chi^2(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.46</td>
<td>0.12</td>
<td>-12.05</td>
<td>15.96***</td>
</tr>
<tr>
<td>Image</td>
<td>0.19</td>
<td>0.05</td>
<td>4.00</td>
<td>20.86***</td>
</tr>
<tr>
<td>Verb Type</td>
<td>-0.01</td>
<td>0.13</td>
<td>-0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>Time linear</td>
<td>1.92</td>
<td>0.32</td>
<td>6.01</td>
<td>3.11</td>
</tr>
<tr>
<td>Time quadratic</td>
<td>0.33</td>
<td>0.15</td>
<td>2.22</td>
<td>42.08***</td>
</tr>
<tr>
<td>Image * Verb Type</td>
<td>0.06</td>
<td>0.09</td>
<td>0.65</td>
<td>0.42</td>
</tr>
<tr>
<td>Image * Time linear</td>
<td>1.46</td>
<td>0.23</td>
<td>6.49</td>
<td>21.32***</td>
</tr>
<tr>
<td>Image * Time quadratic</td>
<td>0.16</td>
<td>0.21</td>
<td>0.74</td>
<td>0.55</td>
</tr>
<tr>
<td>Verb Type * Time linear</td>
<td>0.21</td>
<td>0.52</td>
<td>0.40</td>
<td>0.03</td>
</tr>
<tr>
<td>Verb Type * Time quadratic</td>
<td>0.39</td>
<td>0.28</td>
<td>1.39</td>
<td>1.53</td>
</tr>
<tr>
<td>Image * Verb Type * Time linear</td>
<td>2.09</td>
<td>0.45</td>
<td>4.62</td>
<td>11.58***</td>
</tr>
<tr>
<td>Image * Verb Type * Time quadratic</td>
<td>1.44</td>
<td>0.42</td>
<td>3.40</td>
<td>21.48***</td>
</tr>
</tbody>
</table>

Notes. ***$p < .001$

Table 22. Summary of the grow curve regression results for High L2 group

<table>
<thead>
<tr>
<th></th>
<th>Estimate ($\beta$)</th>
<th>SE</th>
<th>$t$</th>
<th>$\chi^2(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.56</td>
<td>0.14</td>
<td>-11.19</td>
<td>34.87***</td>
</tr>
<tr>
<td>Image</td>
<td>0.37</td>
<td>0.06</td>
<td>5.91</td>
<td>18.36***</td>
</tr>
<tr>
<td>Verb Type</td>
<td>0.22</td>
<td>0.15</td>
<td>1.50</td>
<td>0.92</td>
</tr>
<tr>
<td>Time linear</td>
<td>1.55</td>
<td>0.28</td>
<td>5.51</td>
<td>0.00</td>
</tr>
<tr>
<td>Time quadratic</td>
<td>0.23</td>
<td>0.21</td>
<td>1.12</td>
<td>0.00</td>
</tr>
<tr>
<td>Image * Verb Type</td>
<td>0.00</td>
<td>0.13</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Image * Time linear</td>
<td>0.44</td>
<td>0.31</td>
<td>1.43</td>
<td>2.03</td>
</tr>
<tr>
<td>Image * Time quadratic</td>
<td>-0.18</td>
<td>0.28</td>
<td>-0.63</td>
<td>0.39</td>
</tr>
<tr>
<td>Verb Type * Time linear</td>
<td>-0.39</td>
<td>0.46</td>
<td>-0.84</td>
<td>0.79</td>
</tr>
<tr>
<td>Verb Type * Time quadratic</td>
<td>0.15</td>
<td>0.41</td>
<td>0.37</td>
<td>0.20</td>
</tr>
<tr>
<td>Image * Verb Type * Time linear</td>
<td>0.48</td>
<td>0.62</td>
<td>0.78</td>
<td>0.62</td>
</tr>
<tr>
<td>Image * Verb Type * Time quadratic</td>
<td>1.09</td>
<td>0.57</td>
<td>1.90</td>
<td>3.63†</td>
</tr>
</tbody>
</table>

Notes. ***$p < .001$; †$p = .06$
Table 23. **Summary of the grow curve regression results for Intermediate L2 group**

<table>
<thead>
<tr>
<th></th>
<th>Estimate ($\beta$)</th>
<th>SE</th>
<th>$t$</th>
<th>$\chi^2(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.57</td>
<td>0.14</td>
<td>-10.93</td>
<td></td>
</tr>
<tr>
<td>Image</td>
<td>0.46</td>
<td>0.07</td>
<td>7.10</td>
<td>50.34***</td>
</tr>
<tr>
<td>Verb Type</td>
<td>0.09</td>
<td>0.13</td>
<td>0.69</td>
<td>0.55</td>
</tr>
<tr>
<td>Time linear</td>
<td>1.49</td>
<td>0.24</td>
<td>6.22</td>
<td>21.84***</td>
</tr>
<tr>
<td>Time quadratic</td>
<td>0.05</td>
<td>0.19</td>
<td>0.25</td>
<td>0.03</td>
</tr>
<tr>
<td>Image * Verb Type</td>
<td>0.31</td>
<td>0.13</td>
<td>2.37</td>
<td>5.64*</td>
</tr>
<tr>
<td>Image * Time linear</td>
<td>1.95</td>
<td>0.32</td>
<td>6.14</td>
<td>37.66***</td>
</tr>
<tr>
<td>Image* Time quadratic</td>
<td>0.36</td>
<td>0.29</td>
<td>1.23</td>
<td>1.51</td>
</tr>
<tr>
<td>Verb Type * Time linear</td>
<td>0.75</td>
<td>0.37</td>
<td>2.02</td>
<td>3.74</td>
</tr>
<tr>
<td>Verb Type * Time quadratic</td>
<td>-0.09</td>
<td>0.31</td>
<td>-0.29</td>
<td>0.06</td>
</tr>
<tr>
<td>Image * Verb Type * Time linear</td>
<td>1.87</td>
<td>0.64</td>
<td>2.90</td>
<td>8.43**</td>
</tr>
<tr>
<td>Image* Verb Type *Time quadratic</td>
<td>0.86</td>
<td>0.59</td>
<td>1.46</td>
<td>2.13</td>
</tr>
</tbody>
</table>

**Notes.** *p < .05; **p < 0.1; ***p < .001

Table 24. **Summary of the grow curve regression results for Low L2 group**

<table>
<thead>
<tr>
<th></th>
<th>Estimate ($\beta$)</th>
<th>SE</th>
<th>$t$</th>
<th>$\chi^2(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.41</td>
<td>0.12</td>
<td>-11.33</td>
<td></td>
</tr>
<tr>
<td>Image</td>
<td>0.38</td>
<td>0.07</td>
<td>5.78</td>
<td>33.18***</td>
</tr>
<tr>
<td>Verb Type</td>
<td>0.16</td>
<td>0.14</td>
<td>1.15</td>
<td>1.41</td>
</tr>
<tr>
<td>Time linear</td>
<td>1.51</td>
<td>0.25</td>
<td>6.00</td>
<td>93.59***</td>
</tr>
<tr>
<td>Time quadratic</td>
<td>-0.05</td>
<td>0.20</td>
<td>-0.28</td>
<td>0.07</td>
</tr>
<tr>
<td>Image * Verb Type</td>
<td>0.14</td>
<td>0.13</td>
<td>1.07</td>
<td>1.15</td>
</tr>
<tr>
<td>Image * Time linear</td>
<td>1.08</td>
<td>0.32</td>
<td>3.39</td>
<td>11.44***</td>
</tr>
<tr>
<td>Image* Time quadratic</td>
<td>0.12</td>
<td>0.30</td>
<td>0.40</td>
<td>0.16</td>
</tr>
<tr>
<td>Verb Type * Time linear</td>
<td>0.45</td>
<td>0.39</td>
<td>1.15</td>
<td>3.18</td>
</tr>
<tr>
<td>Verb Type * Time quadratic</td>
<td>0.34</td>
<td>0.35</td>
<td>0.98</td>
<td>1.80</td>
</tr>
<tr>
<td>Image * Verb Type * Time linear</td>
<td>3.01</td>
<td>0.64</td>
<td>4.71</td>
<td>22.00***</td>
</tr>
<tr>
<td>Image* Verb Type *Time quadratic</td>
<td>0.19</td>
<td>0.59</td>
<td>0.33</td>
<td>0.11</td>
</tr>
</tbody>
</table>
REFERENCES
http://doi.org/10.1198/tech.2003.s28


http://doi.org/10.1111/cogs.12018


http://doi.org/10.1016/j.cognition.2012.01.002

http://doi.org/10.1111/j.1551-6709.2009.01055.x


http://doi.org/10.1002/9780470996591.ch63
http://doi.org/10.1007/978-81-322-2443-3


increased attentional blink effect: evidence that the difference between bilinguals and monolinguals generalizes to different levels of second language proficiency. *Psychological Research, 77*(6), 728–737. http://doi.org/10.1007/s00426-012-0466-4


