ABSTRACT

Title of dissertation: VERB LEARNING UNDER GUIDANCE
Angela Xiaoxue He, Doctor of Philosophy, 2015

Dissertation directed by: Professor Jeffrey Lidz
Department of Linguistics

Any kind of uninstructed learning, faced by the challenge that any finite experience is consistent with infinitely many hypotheses, must proceed under guidance. This dissertation investigates guided vocabulary acquisition with a focus on verb learning. In particular, it examines some proposed early expectations that the young language learner may hold as guidance in learning novel verbs, and investigates the nature of these expectations from different angles. Four lines of studies are reported, each discussing a different question. Study 1 focuses on the expectation that the grammatical category *verb* picks out the conceptual category *event* – the verb-event bias, and examines the early developmental trajectory of this bias, which may shed light on its origin: whether it is specified within UG or generalized inductively from input. Study 2 further asks how specific/general the learner's initial expectations about verb meanings are, and thus what is the expected degree of extendibility of verb meanings. Study 3 investigates the proposed expectation that the number of event participants aligns with the number of syntactic arguments – the participant-argument-match (PAM) bias, and questions the utility of this bias in face of potential *mismatch* cases; in particular, some plausible 3-participant events are naturally described by

2-argument sentences. Study 4 looks at the proposed expectation that objects name patients (ONP) and asks a question about its exact nature in face of cross-linguistic variation – whether objects are expected to name patients of the *clause's event*, or to name patients of the *verb's event*, and whether it varies cross-linguistically. Together, this dissertation provides new evidence that the language learner acquires verb meanings under guidance, asks new questions about the natures of some verb-learning guides, and highlights several issues the current acquisition theory needs to address.

VERB LEARNING UNDER GUIDANCE

by

Xiaoxue He

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor Jeffrey Lidz, Chair
Professor Valentine Hacquard
Professor Colin Phillips
Professor Alexander Williams
Professor Rochelle Newman, Dean's Representative

PREFACE

All of the work reported in this dissertation was conducted under the direction of Professor Jeffrey Lidz. The work reported in Chapter 4 and Chapter 5 was also under the direction of Professor Alexander Williams. Chapter 4 reports research that was jointly conducted with my colleague Alexis Wellwood.

DEDICATION

To my parents and my husband Da

ACKNOWLEDGEMENTS

There is no best way to express in plain words, in such limited space, my deepest gratitude to everyone who helped me to get to this point, and to everything that shaped me into who I am today. Despite the challenge, here is my attempt.

First and foremost, I would like to thank Jeffrey Lidz, my advisor, teacher, and the most important person in my academic life so far. To me, he is much more than just the person who supervises my research, reads my papers, and gives me tons of valuable feedback; I will remember him as the person who always had my back at moments of struggling and frustration, and who gave me high fives at moments of joy and excitement. He brought me into the wonderful field of acquisition research, facilitated and witnessed my growth from an academic novice to a competent independent researcher, and supported me at every turn of this challenging journey. For all these, I am forever grateful.

Alexander Williams, not officially my advisor, nonetheless advises a great amount of my work. It was under his direction that many hard and abstract ideas got concretized and transformed into interesting projects. His early experiences in China, my home country, and his rich knowledge about Chinese culture, resulted in us sharing many common interests - in language, food, history, people, etc. He is a strict teacher, but I feel very close to him. Besides Jeff, he is the person who guided and supported me all the way through to get here, and to whom I owe my deepest gratitude.

Jeff and Alexander have been my role models as great researchers. I hope someday to develop Jeff's 'magic' of explaining any very complex idea in terms that are easy to understand, and his sharp ability to quickly grasp the gist of things and share his targeted comments, both of which require a deep understanding of the present and related issues. I also hope to become a researcher like Alexander, with sagacious insights and philosophical ways of thinking, as well as a rich knowledge base that inspires and enlightens new ideas. I thank both of them for being the beacons of my academic journey.

Together with Jeff and Alexander, I would like to also thank Valentine Hacquard, Colin Phillips, and Rochelle Newman for serving in my dissertation committee. Thank you all for your effort and time, for your support, and especially for your great questions that helped me think hard and improve my dissertation. Special thanks go to Valentine, who also served in the committee of my candidacy defense and offered fantastic support and useful advice. The candidacy defense might almost have been an even more important milestone to me than the dissertation, considering how particularly challenging the first few years of graduate life had been for me. Thank you, Valentine, for being there during my most difficult time.

All these people – Jeff, Alexander, Valentine, Colin, and Rochelle, I won't remember them as people who 'grilled' me at defenses. Rather, I will remember them as my teachers, who have passed down their knowledge and shared their expertise in courses, talks, discussions and casual chats. Together with them, I would like to take this opportunity to thank all of my teachers at Maryland: Howard Lasnik, Norbert Hornstein, Juan Uriagereka, Paul Pietroski, Naomi Feldman, Bob Slevc (Psychology), Robert DeKeyser (Second Language Acquisition), Yi-Ting Huang (Hearing and Speech), Jonathan Beier (Psychology), Sandra Cerrai (Mathematics), and Yuan Liao (Mathematics). Without their dedicated teaching, I wouldn't have built up my knowledge base of syntax and semantics, the most essential knowledge for a linguist; I wouldn't have appreciated the beauty and richness of cognitive science that embraces so many disciplines, including, most relevantly to me, psychology, philosophy, and computer science; I wouldn't have obtained the skills necessary for me to conduct research projects, like statistics and programming; and I wouldn't have realized how unbounded learning can be, and worth a life-time of dedication.

Many teachers of mine are outside my home department. I got to know them, to learn from them and to work with them, thanks to the fabulous interdisciplinary community of language science at Maryland. I would like to thank Colin Phillips, who supports and facilitates extensive cross-departmental research here. If I were to write a caption for a picture of him, I would choose 'Phillips: Get out of your comfort zone, and you will shine'. I extended myself into other areas and became interested in them by his encouragement, including bilingual language production—thanks go to Bob Slevc, and second language acquisition—thanks go to Robert DeKeyser. I would like to thank Bob Slevc for his continuing support on our project studying bilingual lexical selection and inhibitory control, despite me being an amateur.

Much is learned from professors, but I have learned even more from my peers. To Kenshi Funakoshi and Sayaka Goto, thank you for all the beautiful syntactic trees and derivations to help me understand syntax questions. To Alexis Wellwood, thank you for all the fun discussions on events and meanings, for all the stimuli-making and data-sharing moments, the kind corrections of my errors in using English, and for all your help with my questions on academia, which I always wanted to ask you before asking Jeff. To Sol Lago, thank you for the company of taking daunting statistics courses in the Mathematics department, and for cracking those tricky math problems together. Together with Kenshi, Sayaka, Alexis, and Sol, thanks also go to Dan Parker, Darryl McAdams, Yakov Kronrod, and Megan Sutton, for being a supportive and friendly first-year cohort. Apart from them, I would also like to thank other fellow students within and outside the Linguistics department, for either direct interactions and discussions with me, or indirectly helping to create a superb intellectual environment.

Another group of people I owe my sincere thanks to are those involved with the infant lab (Project on Children's Language Learning). To Tara Mease, the lab manager, thank you for all your hard work that ensured smooth progress on my projects. To my previous undergraduate assistants—Katrina Connell, Tania Delgado, Radhika Lakhani, Stacey Maresco, Ilanna Newman, Tori Peck, and

Sandy Wan, thank you all for your help with subject recruiting, experiment running, and/or stimuli making. I came to the lab with little administrative experience; after years working here, with many things learned from all of you, I am now able to manage multiple projects as a project leader, to coordinate different human resources, to tell a parent what a project is about in a simple and easy way, and to handle all kinds of unexpected problems that may happen in the lab. Thanks, the infant lab and its people, for shaping me into a capable researcher and experimenter.

My appreciations also go to Kim Kwok, Kathi Faukingham and Csilla Kajtar for their help with all kinds of paperwork, travel funds, tax return, foreign student status, etc. Your help has made my life much easier!

I would like to also express my appreciation to Tonia Bleam. She may not remember because this was years ago, but I will never forget her warmth and kindness when I asked her for help about some personal matter in a really desperate situation. Thank you, Tonia!

I also owe many thanks to many people outside Maryland. First, I would like to place on record my gratitude for people of Department of Linguistics at Chinese University of Hong Kong: To Gladys Tang, the then-chair, thank you for exceptionally admitting me into Linguistics as a former major in Economics, opening the door of this wonderful world to me; to my undergraduate teachers Thomas Hun-Tak Lee and Virginia Choi-Yin Yip, thank you for offering me the precious opportunities of doing linguistic research in your labs. Second, I would like to thank several people outside Maryland for having enlightened me with lengthened discussions and/or meticulous feedback, and/or for having helped me with kind advices on my career—thanks go to Sudha Arunachalam, Alex de Carvalho, Anne Christophe, Cynthia Fisher, Cynthia Lukyanenko, Letitia Naigles, Esther Yi Su, and Sandra Waxman.

Besides people from work, my dear friends have been a constant source of support, comfort and fun. I wouldn't have survived these tough years without all of you being there. To Alexis Wellwood, my office-mate for five years, one of my best friends: I miss all the mornings and afternoons in our office—moments when we each focused on our own work, with tapping sounds of the keyboards and sometimes background music that we both loved, and moments when one of us broke the work focus with some joke, some complaint, some question, or some bursting laugh. I miss all the leisure times we spent together, drinking, board gaming, dining, double dating, and all the inspiring discussions we had on a whole range of non-academic topics. I want more of these! To Sayaka Goto, my crochet teacher, my baking tutor, my Japanese cuisine teacher, my one-year roommate, my dear friend: all the joyful moments we shared - cherry blossom, Delaware beach, Billy Goat trail hiking, etc. – are still quite fresh in my memory, and I can't believe you are now across the Pacific Ocean! I miss you. To Kenshi Funakoshi, my dear friend, a man of few words but many actions, I also have few words for you, but I'm certain my gratitude will reach you: thank you for all the silent support all these years, thank you for being there. To Pauline Bao, Joshua

Davis, Zhenzhen Jia, Chuchu Li, Tongyun Li, Ran Liu, Shota Momma, Naho Orita, Carolina Peterson, Yi Ren, Yi Ren (2 different people), Jiao Wang, Kent Wills, Danyan Zha, Yin Zhang, and many many others: you may not know each other, but you are all connected to me as my dearest friends, who have been at least one of the following: sweet companions, firm supporters, patient complaint takers, joy sharers, hangover bodyguards, embarrassing secret keepers, etc.; some of you may not even understand the title of my dissertation, but partial credits for its completion definitely go to you!

My parents, both researchers in History, did not understand why I chose to engage in a field that is, in their words, so 'uncommon'. Despite this, they are perhaps among the very few people who are even more excited than I am about the completion of this work. They deserve more thanks than I know how to give, for having raised me to pursue the best one can achieve.

Da, my classmate from high school, my soul mate, and my husband, of all people, is the only one who never had event a tiny bit doubt about me, even at moments I deeply doubted myself. He knows me better than I know myself; he learns to know things that I love, linguistics included; he is friend of almost all my friends. He is the one who deserves most of my gratitude, but also the one who does not need it.

Everything in my life, every choice I made, every turn I took, better or worse, success or failure, has all contributed to shaping me into who I am today. I am grateful for all of it.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Learning is, given a set of data, figuring out what is the system/machinery that generates the data. Learning a language is to identify the grammar – a finite system with infinite expressive capacity – that generates the observed input. The child language learner is exposed to a finite set of input data, which is consistent with infinitely many hypotheses about the target grammar. Here the child learner faces a challenge: to identify the target grammar, she has to decide which hypothesis, among unboundedly many, is the correct one, but the input (linguistic and extra-linguistic) does not provide sufficient information to tease apart competing hypotheses (i.e. Poverty of Stimulus, Chomsky, 1980). Nevertheless, the learner successfully acquires a native language with a fully developed target grammar. Learnability theorists, therefore, must explain how the above-mentioned challenge is overcome despite indeterminate input.

To overcome this challenge, the learner must bring with her some biased expectations as guidance in language acquisition. These biased expectations may serve as constraints on the space of hypotheses about the target grammar, or as guidance in how to search the space. This dissertation explores the role of guidance in language acquisition by looking at word learning – an aspect of language that was most easily misunderstood as a process that did not need guidance, because observation of the input alone was once thought to be sufficient. In particular, this dissertation investigates the case of *verb learning* – the class of words that is seen acquired relatively late and to which observational learning is shown to be particularly insufficient.

1.1 Learning Must Be Guided

The language learner's task is to figure what grammar generates the observed input. To achieve this, at least two important steps must be taken. First, the learner has to parse and encode the input into meaningful representations, that is, to build an intake; for example, segmenting individual words out of the continuous speech stream, building a syntactic structure, categorizing words into grammatical categories, building some kind of event representation of the stretch of world concurrent with speech, ec cetera. Second, the learner needs to make inferences about the grammatical features responsible for the intake representations, that is, build a developing grammar; for example, inferring from non-imperative sentences occurring with no arguments that the target grammar has the argument-drop feature, inferring from wh-sentences with wh-words at the top of sentence that the target grammar has the wh-movement feature, ec cetera. These two steps lead to a developing grammar that includes all the inferred grammatical features. Meanwhile, these grammatical features (i.e. the developing grammar) keeps feeding back to these two processes with updated information. See Figure 1.1 for an illustration. For an elaborated discussion of the language acquisition model alluded here, see Lidz and Gagliardi (2014).

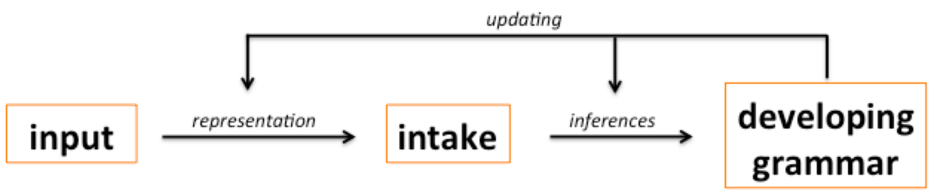


Figure 1.1: Important steps in acquiring a grammar

Now, zoom in on the particular aspect of language acquisition of interest – vocabulary acquisition. To learn the meaning of a word is to form a link between the phonological form of the word and the concept it picks out. Suppose the learner hears the sentence 'Joe frightened the cat' and sees a scenario as shown in the picture[1] on the bottom-left corner of Figure 1.2. An intake representation from the linguistic input will be built: the word form [fɹaɪtən] belongs to the grammatical category *verb*, taking two argument NPs (NP_1 = 'Joe', NP_2 = 'cat'). Meanwhile, the learner will be building some intake representations from the extralinguistic input. Assuming she is attending to the relevant part of the world under discussion, there are still lots of concepts under which this stretch of world can be viewed; in fact, not only lots, but infinitely many; Figure 1.2 shows just a few examples. From these intake representations, the learner will then need to infer what is the target concept this word form [fɹaɪtən] maps onto, and then add

[1] Image retrieved from http://www.anvari.org/cols/Amazing_Moments.html on Mar. 30, 2015

that piece of information into the developing lexicon. See Figure 1.2 for an illustration of this process.

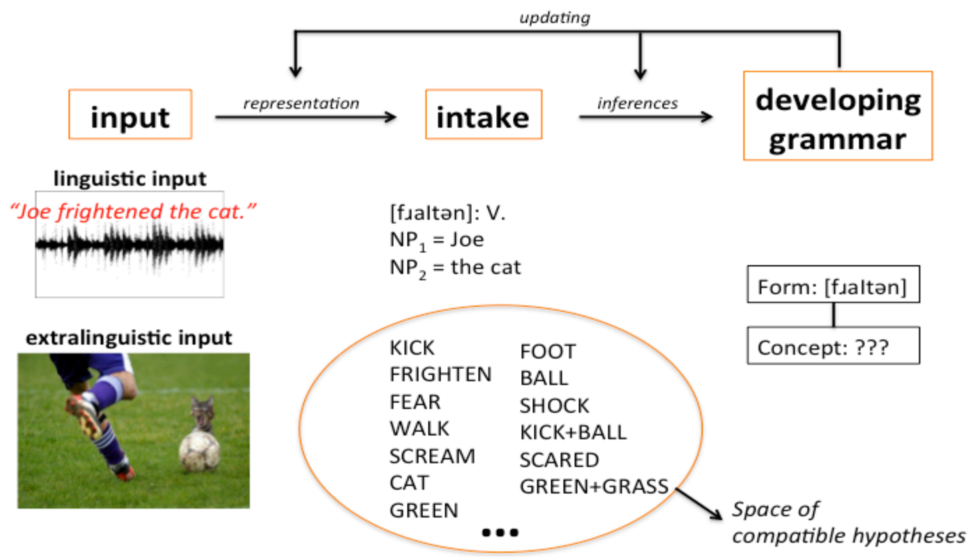


Figure 1.2: A step-by-step word-learning example

Vocabulary acquisition was once considered in history a straightforward process of establishing a mapping between the word form and the relevant part of the world by observing the contingencies of word use (cf. Locke, 1690/1959), known as 'word-to-world mapping'. This standpoint is based on the assumption that observation of the input alone is sufficient for the learner to identify which part of the world is relevant, and furthermore, under which concept that relevant part of world is viewed. The Poverty of Stimulus argument would be hardly raised for word learning if this assumption were valid. It has become widely recognized nowadays, however, that this assumption is not valid, because the context of use (i.e. the world) makes available infinitely many candidate concepts for a word form to map onto, which are not distinguishable by the input alone. In the above case, for instance, the word *frighten*, unbeknownst to the child learner, could in principle pick out any aspect of the world, including an object at the basic (e.g. CAT) or superordinate (e.g. ANIMAL) level, some other object (e.g. BALL, GRASS), a part of an object (e.g. HEAD), a property of an object (e.g. GREEN), a status of an object (e.g. SCARED), an action (e.g. KICK, RUN), a psychological event concept (e.g. FRIGHTEN, FEAR), some combination of these features (e.g. KICK + BALL, GREEN + GRASS, CAT + FEAR), or even something irrelevant to the present scene (IT'S SUNNY TODAY), and so on so forth[2]. These are all possible hypotheses about the meaning of the target word, which are indistinguishable by the input itself. Therefore, vocabulary acquisition, as much as acquisition of syntax, presents a classic Poverty of Stimulus problem (Chomsky, 1959; Gleitman, 1990; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Lidz, 2006; Quine, 1960).

To overcome this problem, the process of word learning needs to be properly guided. Guidance may come in the form of constraints on the space of possible meanings that the learner considers for a novel word or as directions in how to search the space. About the necessity of having guidance, Gleitman (1990) wrote as follows:

[2] Words in capital letters stand for the concepts/meanings picked out by the corresponding word forms, which are italicized.

> "The trouble is that an observer who notices *everything* can learn *nothing*, for there is no end of categories known and constructible to describe a situation. Indeed, not only learnability theorists but all syntacticians in the generative tradition appeal to the desirability of narrowing the hypothesis space lest the child be so overwhelmed with representations options and data-manipulative capacity as to be lost in thought forever." (p. 12)

Guidance is necessary for learning any category of words (e.g. nouns, adjectives and verbs). Count nouns, with which observable objects are regularly labeled, still have too large a space of possible meanings if unconstrained. For example, any time the object – a cat – occurs, its parts, its superordinate level concept, its properties, its state or action, etc., all occur with it. The problem is not lessened by the referents being observable. Constraints/directions guiding noun learning may include the following: the expectation that novel count nouns refer to object kinds rather than properties of or relations between objects (i.e. the Noun-Object link; Gentner, 1982; Waxman & Booth, 2001), the expectation that novel count nouns refer to whole objects rather than object parts (i.e. the Whole Object bias; Markman & Hutchinson, 1984; Soja, Carey, & Spelke, 1991), and the expectation that novel nouns refer to the basic-level kinds rather than superordinate-level kinds (Markman, 1991; Mervis, 1987). These constraints may provide guidance in early noun learning, leading to a steady growth of count noun vocabulary from the learner's first birthday to about the 17^{th} to 18^{th} month (for a review of noun learning constraints, see Woodward, Golinkoff, Hirsh-Pasek, & Bloom, 2000).

Acquisition of verbs, too, perhaps to a greater extent, must proceed under guidance. Compared to nouns, acquisition of verbs is delayed - it is not until the second birthday that children start to produce a sizeable number of verbs and use them systematically (Bates et al., 1994; Caselli et al., 1995; Fenson et al., 1994; Gentner, 1982). Verb meanings are also shown to be particularly hard to learn from observation of extra-linguistic context alone. In a task that uses adult participants to simulate word learning, Gillette, Gleitman, Gleitman, and Lederer (1999) showed that adults correctly guessed nouns three times more often than they correctly identified verbs from observing scenes of parents talking to their child (although correct guesses were not very common even for nouns). This dissertation focuses on this class of late acquired words that are particularly recalcitrant to observational learning to investigate vocabulary acquisition under guidance. In particular, it examines some proposed *early expectations* the learner may hold to guide her inferences about possible verb meanings, and investigates the nature of these expectations from different aspects; see Figure 1.3. In the following section, I discuss four such issues.

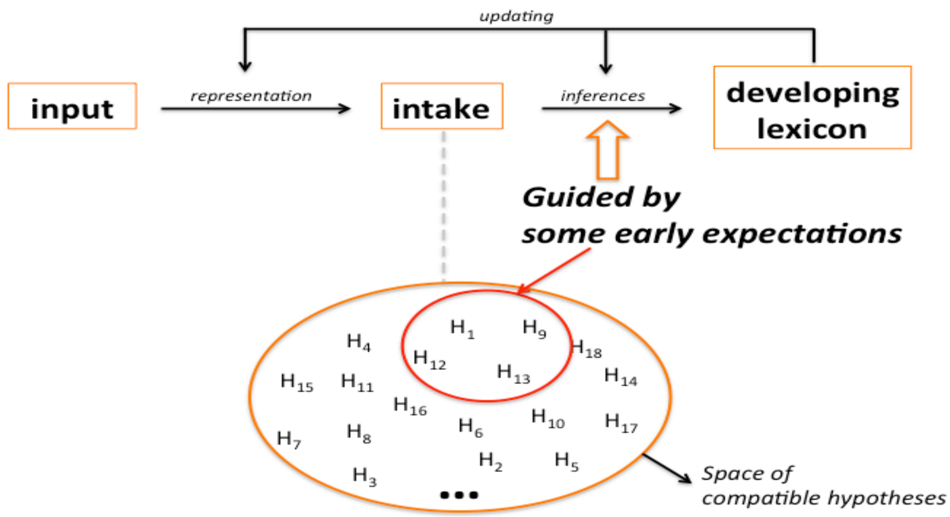


Figure 1.3: Building a developing lexicon under guidance

1.2 Some Early Expectations Guiding Verb Learning

1.2.1 Expecting Verbs to Pick Out Event Concepts

One possible guide in verb learning is the expectation that the grammatical category *verb* picks out the conceptual category *event*. Call this the *verb-event bias*.

This proposal derives from the observation that different types of words tend to pick out different aspects of the world (Waxman & Booth, 2001). Here, *types of word* mean the formally defined syntactic categories[3] (e.g. nouns, adjectives, verbs), and *aspects of the world* mean categories of concepts (e.g. objects, properties, events) under which the world is viewed. For example,

[3] This dissertation uses 'syntactic category' and 'grammatical category' interchangeably.

predicates that stand for sorts of objects are almost always nouns (Greenberg, 1963; Macnamara, 1982); modality, tense, aspect, and epistemic status are usually designated by auxiliaries (Steele, 1981); verbs often depict events, states and relations. Such correspondences are by no means exclusive or deterministic; for instance, some event concepts (e.g. EARTHQUAKE, REHEARSAL) are expressed by nouns (e.g. *earthquake*, *rehearsal*). But they certainly capture the probabilistic correlations between syntactic categories and conceptual categories. These correlations are so strong that lay people usually talk about nouns as substance/object words and verbs as action/relation words (Brown, 1957), and that linguists use conceptual notions as their first hypothesis about the syntactic category of an unknown word and such a non-formal strategy seldom leads to trouble (Pinker, 1984b, personal communication with Jane Grimshaw). Importantly, such correspondences seem even stronger in children's speech than in adults': in a study of nouns and verbs in children's speech, Brown (1957) found 67% of children's use of nouns had visual contour (16% in adults' nouns) and 83% had size implication (39% in adults' nouns), both of which were defining features of picturability and thus concreteness; and he found 67% of children's verbs named movements (33% in adults' verbs). Brown (1957) raised the possibility that a set of "firm, temporarily reliable" relations between certain syntactic categories (like nouns and verbs) and semantic notions (like concrete objects and actions) develops in children, and stays till adulthood but only retains as "a probabilistic truth" (pp. 2). He further proposed that these temporarily reliable relations may allow the syntactic categories to operate as a filter for selecting candidates concepts (pp. 3).

Based on the correlational structure between syntactic categories and conceptual categories, in particular between *verb* and *event*, it is hypothesized that the verb-event correlation is known to the learner, and facilitates verb learning in the following way: the learner categorizes a novel word into the syntactic category *verb*; fixing this source information activates the inference chain from *verb* to *event*, thus narrowing the learner's hypothesis space down to (or directing the leaner's search to) event concepts (saving the effort to consider concepts of object kinds or property types). Research has shown deployment of this verb-event link in novel verb learning tasks in infancy (i.e. around the end of second year), confirming the part of the proposal that learners do have access to this correlational link, expecting verbs to pick out event concepts (as opposed to other categories of concepts). But the origins of this expectation remains unclear: whether it is part of the Universal Grammar, pre-programed into the learner's language acquisition device, or, alternatively, generalized from experience via inductive analysis. In this respect, an examination of the developmental trajectory of the verb-event bias may shed some light on disentangling these two sources. If the bias is part of UG, then a timetable for its establishment is pre-programed; specifically, it should be activated once all the representations supporting its operation are established – that is, once the learner discovers the surface expressions of the *verb* category in the particular language she is learning, then the expectation that words in this category pick out event concepts should follow for free. Further, if it is specified by UG as a learning bias to facilitate acquisition of some particular piece of knowledge, it should be deployed by the learner prior to the piece of knowledge being seen acquired. Here, it may be worth pointing out that being part of UG does not imply emergence at birth (a common misconception), in a similar way to that teeth growth is part of human biology but is not seen at birth. If the bias is generalized from experience, on the other hand, it is more likely to emerge slowly, through comparison across many individual instances learned one at a time (cf. Tomasello, 2000).

Chapter 2 of this dissertation examines the early developmental trajectory of the verb-event bias, which may shed some light on its origin: whether the growth of a verb vocabulary supports the induction of the verb-event link or whether the activation of the verb-event link supports rapid growth in verb vocabulary. To preview, experimental data from this work provides evidence that this bias is not the result of abstracting a common semantic core out of a large set of already acquired verbs, but rather provides the scaffolding for the growth of the verb vocabulary.

1.2.2 Expecting Specific versus General Verb Meanings

The learner may expect verbs to pick out event concepts, but it remains unclear how specific her initial expectations about verb meanings are. Some verbs place very specific restrictions on their arguments; for example, one can only *dribble* something that can bounce, one can only *eat* things that are edible, only animals can *hibernate*, ec cetera; these verbs are narrowly extendible. Other verbs are quite general; for example, anything can *fall* or *spin*, anything can be *dropped* or *kicked*, ec cetera; these verbs are broadly extendible. A complete mastery of verb meanings include a) ability to generalize the verb beyond the particular exemplar in which it is first learned, and to extend it to other events of the same type, and b) knowledge of the specific semantic restrictions the verb places on its arguments.

This dissertation is not committed to address the question about when and how specific semantic restrictions are acquired (which is equally fascinating and important), but only concerns the initial expectations the learner may hold in this respect. In particular, it asks to what extent the learner's initial expectations of verb meanings are specific versus general. For instance, if the learner first learns the verb *run* from the instance of a cat running, is she biased to expect the event concept picked out by *run* to be the category of running events of cat sorts of thing (i.e. more specific), or the category of running events of any animate objects (i.e. more general)? The extent of specificity/generality to which the learner's initial expectations of verb meanings are determined, and therefore can be tested through, the extent to which the learner is willing to extend a newly-learned verb to new instances of events of the same type but with different sorts of participants. In the same example as above, the more broad expectation (i.e. *run* picks out running type of events of any animate objects) may lead the learner to readily accept an instance of a horse running as an exemplar concept of the verb *run*; she may not be willing to accept so, however, if the more specific expectation (i.e. *run* picks out running type of event of cat sorts of thing) is held.

Mixed findings are documented in the literature, some showing an early expectation of at least some degree of generality evidenced by ready extension of verbs to same event types with different sorts of participants, whereas some showing very specific expectation – one that ties verb meanings to some particular event participants, evidenced by failure of verb extension prolonged to preschool years. Chapter 3 of this dissertation examines this issue; in particular it evaluates the possibility that verb extension failures in previous work may be an artifact of task requirements rather than reflecting true grammatical knowledge, that is, deployment of certain knowledge may sometimes be obscured by performance limitation. To preview, an experiment that took special efforts in minimizing cognitive burdens revealed young infants' willingness to extend newly learned verbs to new instances with different sorts of participants, lending more support for the view that the learner expects verb meanings to be relatively general and thus relatively broadly extendible.

1.2.3 Expecting Participant Number to Match Argument Number

Verbs pick out event concepts. For example, the verb 'cut' picks out a cutting event. Verbs also occur in certain syntactic environments. For example, the verb *cut* occurs in transitive clauses with two noun phrases, one denoting the agent of the cutting event, and the other the patient. Another early expectation the learner may bring with her to facilitate verb learning, as proposed in a prominent theory – the Syntactic Bootstrapping Hypothesis (Fisher, Gertner, Scott, & Yuan, 2010;

Gleitman, 1990; Landau & Gleitman, 1985) – is to use syntax as a 'mental zoom lens' to fix on the target meaning (Gleitman, 1990).

> "Children's sophisticated perceptual and conceptual capacities yield a good many possibilities for interpreting any scene, but the syntax acts as a kind of mental zoom lens for fixing on just the interpretation, among these possible ones, that the speaker is expressing." (pp. 23)

The fundamental idea of this hypothesis is: there are certain principled relations between verb meaning and verb syntax, which is known to the child learner, and thus the learner can use this knowledge to infer verb meaning from verb syntax. This principled mapping between verb meaning and verb syntax, in its strongest version, states that the semantic space of verb meanings is partitioned based on the range of subcategorization frames (Gleitman, 1990). The mappings are by no means deterministic, but are merely correlational, just as the correlations between grammatical categories and conceptual categories (Section 1.2.1). What makes such correlations good candidates for early learning biases is that they have been shown to be quite strong and reliable: Fisher, Gleitman, and Gleitman (1991) showed that for a given set of verbs, the semantic partitioning obtained from adults' judgments of semantic relatedness of the verbs (by choosing the semantic outlier among triads), and the syntactic partitioning obtained from judgments of grammaticality of the verbs in all kind of syntactic environments, had a large degree of overlap.

One important aspect of this theory, as fleshed out in Lidz, Gleitman, and Gleitman (2003) and Lidz and Gleitman (2004), captures the relation between the number of event participants and syntactic arguments: "Every participant in an event as it is mentally represented shows up as a syntactic phrase in a sentence describing that event" (Lidz & Gleitman, 2004, pp. 1). For example, the verb *swim* requires one noun phrase as its argument (n = 1), as in 'Joe swims', and the event it picks out plausibly has one participant (m = 1) – presumably the swimmer; similarly, pound plausibly picks out an event with two participants (m = 2) – the pounder and the thing being pounded, and shows up in sentences with two arguments (n = 2), as in 'Joe pounded the table'; and the n = 3 case is illustrated by give, as in 'Joe gave a book to Jane', which plausibly picks out an event with three participants (n = 3) – the giver, the recipient and the thing given[4]. Given such correlations, learners could use the number of phrasal arguments in the sentence to infer the number of participants in the event denoted as it is mentally represented, so as to narrow down their space of hypotheses (or to direct their search) about the meaning of the verb, i.e. what type of event the verb picks out. This could serve as a second-pass filter about possible verbs meanings on top of the verb-event bias. In the rest of the dissertation, I will be referring to this aspect of Syntactic Bootstrapping theory as the *Participant-Argument-Match* hypothesis, or *PAM* for short.

PAM states that a sentence with n arguments directs the learner's attention to an event with m participants (m = n), among multiple available events. To test

[4] *n* denotes the number of arguments, *m* the number of participants; specific arguments are underlined

this hypothesis, in previous work, children were usually shown two events – a 1-pariticpant one (presumably so) and a 2-pariticpant one (presumably so), and they either heard an intransitive sentence (1-argument) or a transitive sentence (2-argument). Results obtained from most previous work showed that children's attention to either of the events was modulated by the type of sentence they heard. This has been taken as supporting evidence for this hypothesis, which is fair; but its presumption about how the extralinguistic stimuli (i.e. the candidate events) are viewed by children needs to be justified: How do we know that a child is viewing an event as having m participants ($m = n$), rather than m' participants ($m' \neq n$)? After all, the world itself is not labeled by participant roles, and any stretch of the world can be viewed under many concepts. To develop an adequate test for this hypothesis, Chapter 4 of this dissertation examines event representation experimentally, focusing on a relatively under-studied case, where a 2-argument sentence ($n = 2$) plausibly expresses a 3-participant event concept ($m = 3$) – a potential mismatch in argument and participant number ($m \neq n$). For example, 'Anne robbed Betty' has no evident argument for what was stolen, and 'Anne jimmied the box' has no evident argument for the implied lever. If such potential mismatch turns out to be true – that is, if the event is represented as having three participants but the sentence contains only two arguments, then it poses a real challenge to the hypothesis that the learner uses number match between arguments and participants to guide verb learning. To preview, this dissertation presents experimental data that suggest this challenge is real, and therefore calls for modification or further specification of this hypothesis.

1.2.4 Expecting Objects to Name Patients

While PAM concerns the number correlation between participants and arguments, another aspect of the Syntactic Bootstrapping Hypothesis concerns the correlation between the *syntactic position* of an argument and the *thematic role* the corresponding event participant takes; for example, the subject noun phrase corresponds to experiencer or agent, direct object to patient or theme, and indirect object to paths or goals. This dissertation focuses on the correlation between the *direct object* position and the *patient* role (Dowty, 1989, 1991), and examines a possible expectation the learner may hold: objects name patients. This expectation will be referred to as the *Objects Name Patients (ONP)* expectation. ONP may guide verb learning in this way: for example, hearing a sentence like 'the bunny is *gorping* the duck', guided by this expectation, the learner could infer that the novel verb *gorp* names an event of which the duck is a patient.

In Chapter 5 I introduce a new question about the exact nature of this expectation: does the learner expect the object to be the patient of the event described by the *clause*, or instead, the event of the *verb*? This question is empty when the two events – event of the clause and event of the verb – are the same, as they may be when the predicate is simple. For example, the sentence 'Al pounded the cutlet' is made true by events of pounding, and plausibly the verb *pound* itself expresses the concept of being a POUNDING. But it might be that the two events are not always the same, particularly when the predicate is complex. Rather, the clause may describe an event that is *related* to the event of the verb, and not identical to it. Take the sentence 'Al pounded the cutlet flat' for example; here, the event described by the clause is a POUNDING-FLAT event, which is related, but not identical to, the event described by the verb, a POUNDING event. Therefore, to see the exact nature of this ONP expectation, we have to look at the context of complex predicates.

What makes the question even more intriguing is some observed cross-linguistic variation: in some languages (e.g. English), the arguments correspond to the participants picked out by the verb's event – objects name patients of *the event of the verb*, whereas in some languages (e.g. Mandarin), the arguments correspond to the participants picked out by the clause's event – objects name patients of *the event of the clause*. Chapter 5 of this dissertation examines the exact nature of this expectation in English- and Mandarin-learning toddlers, and asks whether young learners are sensitive to the language-specific properties with regard to this expectation. To preview, it provides evidence that young toddlers are indeed sensitive to the specific version of objects-name-patient expectation from very early (i.e. by two-and-a-half years of age). This observation raises new questions for the acquisition theory, in particular, in what ways the current theory could accommodate early expectations of different versions that are language-specific.

1.3 Some Assumptions

To investigate the above four issues, this dissertation operates on the following assumptions: by the time the learner begins to utilize these expectations to guide verb learning, she is able to build appropriate intake representations of the linguistic input; in particular, she is able to represent the grammatical category *verb*, and able to parse sentences into some basic argument structures. These assumptions are important: first, inferences from the grammatical category *verb* to the event category *event* (Section 1.2.1), and deciding how narrowly/broadly extendible the verb is (Section 1.2.2), both require the learner to properly represent a novel word as belonging to the *verb* category; and second, inferences from the number of arguments a verb takes to an event concept with the same number of participants (Section 1.2.3), and inferences from the position of an argument to the thematic role it bears, both require the learner to properly represent the sentence's argument structure. Previous research has lent credibility to these assumptions. Below, I review some evidence.

1.3.1 Representation of Syntactic Category

For inferences from the grammatical category *verb* to the conceptual category *event* to go through, or for inferences about the degree of extendibility of a verb's meaning, it requires a proper representation of the category *verb* to be available.

A well-known proposal states that frequently occurring function words (e.g. determiner, pronoun, auxiliary, etc.) and morphemes (e.g. tense marker, aspect marker, plural marker, etc.) provide distributional cues to the grammatical categories of content words (e.g. noun, adjective, verb). It is shown that infants by one year of age already have some knowledge of these function words and morphemes: for example, Hallé, Durand, and de Boysson-Bardies (2008) demonstrated with the Head-Turn Preference procedure that 11-month-old French-learning infants were able to distinguish real and pseudo articles; and similarly, Shafer, Shucard, Shucard, and Gerken (1998) showed that English-learning 11-month-olds were able to distinguish real English function morphemes from modified atypical ones, indicated by significantly lower amplitude ERPs elicited by the atypical morphemes[5]. Regarding the relevant linguistic underpinnings that support the categorization of *verb*, in particular, we know from previous literature three things.

First, analyses of child-directed speech corpora have identified various morphosyntactic cues to grammatical categories in speech to children (Cartwright & Brent, 1997; Chemla, Mintz, Bernal, & Christophe, 2009; Maratsos & Chalkley, 1980; Mintz, 2003, 2006; Mintz, Newport, & Bever, 2002; Redington, Crater, & Finch, 1998); for example, frequent frames like 'the __ is' or 'you __ it' are good cues to nouns and verbs, respectively (Mintz, 2003).

Second, evidence from behavioral studies has shown that infants are sensitive to these cues; for example, Santelmann and Jusczyk (1998) showed that 18-month-old English-learning infants could distinguish a well-formed morphosyntactic dependency of verbs – the 'is __ing' dependency, from an ungrammatical ('can __ing') dependency.

[5] Also see Shi, Werker, and Cutler (2006) for documentation of a developmental trend of recognition of English common morphemes in young infants.

Third, infants were shown to be able to utilize these cues for novel word categorization (Bernal, Dehaene - Lambertz, Millotte, & Christophe, 2010; Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; Mintz, 2006; Peterson-Hicks, 2006) (also see Shi, 2014, for a review). Evidence for infants' early categorization of *verb* comes from both neurological measures and behavioral studies. Bernal et al. (2010) observed an early left-lateralized (ELAN) brain response[6] in 24-month-old French infants when hearing a noun incorrectly inserted in a verb position (or vice versa). In addition, two behavioral studies using Head-Turn Preference Procedure (Jusczyk & Aslin, 1995) demonstrated verb categorization in younger infants. Peterson-Hicks (2006) showed that 15-month-old infants familiarized with novel words in frequent verb frames (e.g. 'can *pell*', 'will *pell*') listened longer to those words presented in ungrammatical test frames (e.g. 'her *pell*', 'my *pell*') than those presented in grammatical test frames; importantly, the grammatical frames at test were different from the frames used in familiarization (e.g. infants familiarized with 'can *pell*' were tested with 'will *pell*'). This suggested that infants this age were able to use preceding auxiliaries to categorize a novel word into the verb category. Mintz (2006) reported that even younger infants were able to do so: 12-month-olds familiarized with novel words in frequent verb frames (e.g. 'can *gorp*', 'to *gorp*'), listened longer to those words presented in ungrammatical test frames (e.g. 'the gorp', 'a *gorp*') than those

[6] ELAN is an ERP response that most often occurs when linguistic stimuli violate word-category or phrase structure rules.

presented in novel grammatical test frames (e.g. 'to *gorp*' for those who heard 'can *gorp*' during familiarization, and 'can *gorp*' for those who heard 'to *gorp*').[7]

These findings suggest that by the end of the first year at the earliest, the learner may have established a representation of the grammatical category corresponding to what we know as *verb* – a category of words that occur in certain distributions (e.g. 'can __' or 'will __' in English).

1.3.2 Representation of Syntactic Structure

For inferences from *number and positions of arguments* to *number and thematic roles of participants* to go through, it requires a proper representation of the syntactic structure to be available.

It is claimed that a rudimentary syntactic structure of sentence may be derived from a first-pass analysis of the sentence's phonological structure (termed as Phonological Bootstrapping in Morgan & Demuth, 1996)[8], because phrase structures are marked by some phonological correlates (cf. Gleitman & Wanner, 1982; Jusczyk et al., 1992; Morgan, 1986; Morgan & Demuth, 1996). Support for this idea comes from the following findings.

[7] To be fair, Höhle et al. (2004) did not find successful verb categorization (e.g. *sie glummin*, 'she glummin') in 16-month-old German-learning infants, but only found successful noun categorization (e.g. *das Pronk*, 'the pronk').
[8] Similar idea is also termed as *prosodic bootstrapping* in other contexts. Morgan and Demuth (1996) called it *phonological bootstrapping* to emphasize that not only prosodic information, but also phonetic, phonotactic, and stochastic information, all contribute to building syntactic representations.

First, phonological correlates to phrase structures exist in the speech stream. For example, phonological phrase boundaries, characterized by final lengthening, initial strengthening, reduced phoneme coarticulation across the boundary, and a single pitch contour (cf. Christophe, Gout, Peperkamp, & Morgan, 2003; Gout, Christophe, & Morgan, 2004; Shattuck-Hufnagel & Turk, 1996; for reviews), often coincide with syntactic constituent boundaries. For another instance, the location of prominence within a phonological phrase often coincides with the head-direction of the language – for head-initial languages, prominence falls on the right and for head-final languages, prominence falls on the left (Nespor & Vogel, 1986).

Second, infants are shown to be able to perceive these correlates very early on. Christophe, Dupoux, Bertoncini, and Mehler (1994) showed that three-day-old French newborns were able to distinguish bisyllabic strings that contain a phonological phrase boundary (i.e. extracted from within words) from those that do not (i.e. extracted from between words). Hirsh-Pasek et al. (1987), Gerken, Jusczyk, and Mandel (1994), and Jusczyk et al. (1992) demonstrated pre-linguistic infants were sensitive to disruptions (i.e. inserted pauses) of phonological phrases, discriminating samples where pauses were inserted at the boundaries from those where pauses were inserted within the boundaries. Christophe, Nespor, Guasti, and van Ooyen (2003) showed 2- to 3-month-old infants were able to discriminate sentences in a head-initial language (i.e. French) and sentences in a head-final language (i.e. Turkish), based solely on prosodic cues (i.e. prominence within phonological phrases).

Third, it has been shown that young infants are able to exploit phonological phrase boundaries to postulate word boundaries, so as to facilitate word segmentation. Gout et al. (2004) showed that 10- and 13-month-old English-learning infants, were able to recognize a bisyllabic word (e.g. *paper*) from continuous sentences if the two syllables occur within one phonological phrase (e.g. '[the college] [with the biggest *paper* forms] [is best]'[9]), but not able to do so when the two syllables were separated by a phonological phrase boundary (e.g. '[the butler] [with the highest *pay*] [*per*forms the most]'). Similar results were found with English-learning 12-month-olds (Johnson, 2008) and with French-learning 16-month-olds (Millotte, 2005).

Lastly, with regards to whether infants exploit phonological correlates to build more complex syntactic representations, direct evidence is limited. Some indirect evidence comes from studies on adults, showing adults' online syntactic analyses are affected by phonological phrase boundaries. Millotte, Wales, and Christophe (2007) created locally ambiguous sentences (e.g. 'les pommes *dures* …'), which contained homophones that could belong to different syntactic categories (e.g. *dures* could be an adjective meaning HARD, or a verb meaning LAST), and differed in their prosodic structures (e.g. in the adjective case, '[les pommes *dures*] …'; in the verb case, '[les pommes] *dures* …'); French adults, who were presented with these sentences cut before their disambiguation points and asked to complete the sentences, successfully exploited the prosodic cues to disambiguate the homophones, assigning them to their correct syntactic

[9] Brackets are used to indicate phonological phrases.

categories; see Millotte, René, Wales, and Christophe (2008) for a replication of these results in a similar sentence completion task, and more support from an online word detection task.

Taken together, direct and indirect evidence all lend support to the plausibility of the learner arriving at some rudimentary syntactic structure representations via phonological analyses of the input speech.

From there, increasing knowledge of nouns in the learner's early vocabulary at around 17th to 18th month (Caselli et al., 1995; Gentner, 1982) scaffolds the building of an argument structure (argument positions filled by noun phrases) (Golinkoff & Hirsh-Pasek, 2008); this provides the prerequisites for counting the number of arguments. Meanwhile, sensitivity to the basic word order of the target language has been seen by the end of the second year (Gertner, Fisher, & Eisengart, 2006; Hirsh-Pasek, Golinkoff, & Naigles, 1996; Seidl, Hollich, & Jusczyk, 2003); this provides the prerequisites for activating the inference from argument position to thematic role. Therefore, it is reasonable to assume the source information enabling inferences from the syntactic structure to the conceptual structure - a partial syntactic representation - is available to the learner by two years of age (Yuan, Fisher, & Snedeker, 2012).

This dissertation takes as assumptions what previous research has shown about the learner's ability to representing the linguistic input, and from there, it examines the four issues introduced in Section 1.2. Assuming that the syntactic end of the inference chain is fixed and that the direction of inference goes from syntax to meaning is adopting the Syntactic Bootstrapping framework of language acquisition. I must point out, however, that taking this perspective by no means implies an argument against the Semantic Bootstrapping point of view, which assumes the semantic end of the inference chain is fixed and the direction of inference goes from semantics to syntax. I believe a comprehensive theory of language acquisition must embrace both perspectives and admit that they go hand in hand in some complementary manner. This dissertation does not commit itself to the tension/interaction between these two equally enlightening perspectives; instead, it adopts the Syntactic Bootstrapping viewpoint, from which investigations of new questions are underway.

1.4 Summary and Overview

This chapter sets off with a puzzle in acquisition: from sound waves (i.e. linguistic input) and entities and happenings (i.e. extralinguistic input) in the world, how does the child learner successfully acquire a full-blown target grammar that generates the observed input, provided that there are many hypotheses compatible with and yet indistinguishable by the input? This challenge, characterized as the Poverty of Stimulus problem by Chomsky (1980), is omnipresent in every aspect of language acquisition, word learning included. In vocabulary acquisition, any given word form is consistent with infinitely many possible concepts, presenting the learner the challenge of deciding which one is the target concept for the form to map onto. If the learner were going to evaluate an infinite number of hypotheses, the acquisition task would hardly get off the ground. Therefore, it follows that word learning must be guided.

This dissertation investigates guided vocabulary acquisition using *verb learning* as a test case. In particular, it examines some proposed early expectations that the learner may hold as guidance in learning novel verbs. Chapter 2 to 5 will present four studies, each designed to address an issue regarding verb learning under guidance: Chapter 2 examines the proposed expectation that verbs pick out event concepts – the verb-event bias, with the goal to identify the origin of this bias; Chapter 3 asks how specific/general the learner expects verb meanings to be, and thus what is the expected degree of extendibility of verb meanings; Chapter 4 investigates the proposed expectation that the number of event participants aligns with the number of syntactic arguments – the Participant-Argument-Match (PAM) bias, and questions the utility of this bias in face of the observation that some plausible 3-participant events are naturally described by 2-argument sentences; Chapter 5 looks at the proposed expectation that objects name patients (ONP) and asks a question about its exact nature in face of cross-linguistic variation – whether objects are expected to name patients of the clause's event, or to name patients of the verb's event, and whether it varies cross-linguistically. The final chapter, Chapter 6, concludes the dissertation, with a general discussion about the implications of these lines of research on the language acquisition model, and some forward-looking to future research.

Chapter 2[10]

Development of the Verb-Event Bias in Infancy

One possible guidance in verb learning relies on the correspondences between grammatical categories and conceptual categories (Brown, 1957; Grimshaw, 1981; Pinker, 1989b; Waxman & Booth, 2001; Waxman, Lidz, Braun, & Lavin, 2009), and on the inferences made from certain grammatical category to certain conceptual category. For example, if the learner expects *event* categories to be expressed in the grammatical category *verb*, then upon categorizing a word as a *verb*, she will be able to infer with some certainty that this word picks out an *event* concept, avoiding the need to consider meanings from other conceptual categories like *object* or *property*. This constraint, dubbed as *the verb-event bias*, will be the focus of the current chapter.

This chapter investigates the emergence of the verb-event bias at early stages of English-learning infants' lexical acquisition, with the primary goal of mapping the early developmental trajectory of infants' use of this bias in learning novel verbs. This investigation may shed some light on the origin of this bias: if deployment of the verb-event bias is observed prior to a substantial increase in

---

[10] The content in Chapter 2 was also reported in a paper submitted for publication: He, A. X. & J. Lidz (submitted). Development of the verb-event link in 14- and 18-month-old English-learning infants.

verb vocabulary – call this stage the 'verb spurt period', then it suggests the verb-event correspondence supports rapid growth in verb vocabulary, and lends more support for the bias originating from the language acquisition device rather than from experience-based induction; on the other hand, if the deployment of this bias is only seen after a substantial increase in verb vocabulary, or is closely time-locked with the verb spurt period, as documented by previous literature, then it is possible that the verb-event correspondence is induced from a growing verb vocabulary.

In what follows, I will begin with a review of previous literature that motivates this investigation (Section 2.1); then, I will present two experiments designed to test infants' knowledge about the verb-event link, which shows English-learning infants have this knowledge and use it to learn novel verb meanings by 18 months of age - a couple of months prior to the verb spurt period, but provides no evidence for 14-month-olds' knowledge of it (Section 2.2 – 2.4); from there, I will discuss what the findings imply about the role of the verb-event bias in verb learning and the origin of this bias (Section 2.5).

2.1 Background

This chapter aims to map the early developmental trajectory of the learner's deployment of the verb-event bias in learning novel verbs, relative to some other developmental milestones, hoping to shed some light on the origin of this bias. With this goal, this section first reviews previous documentation on first deployment of this bias in verb learning (Section 2.1.1), then discusses why these previous findings provide ambiguous evidence (if not weak) about the origin of this bias, in light of some important developmental milestones relevant to our current discussion, for example, the verb spurt period, and establishment of the prerequisites for activating the verb-event bias (Section 2.1.2), and subsequently, discusses which period in development might be an optimal window for the purpose of the current investigation (Section 2.1.3).

2.1.1 Literature: First Deployment of the Verb-Event Bias

Infants start to produce their first words at the end of the first year, and it has been shown that labeling with a novel noun facilitates object categorization compared to no labeling for infants around the same age, i.e. 12- to 13-months of age (Waxman & Markow, 1995). At 14 months, nouns are treated differently from adjectives in that nouns pick out object categories such as DOG or TABLE, but not object properties such as YELLOW or SQUARE, whereas adjectives highlight both commonalties (Waxman & Booth, 2001), suggesting that the noun-object correspondence emerges soon after the first birthday. This early emergence of the noun-object correspondence may serve as a support for subsequent growth of noun vocabulary: from the first birthday to about the 17^{th} to 18^{th} month, the lexicon grows steadily, with a preponderance of words referring to object categories. It is not until their second birthday that children start to produce a sizeable number of verbs and use them systematically to refer to actions, mental states and relations (Bates et al., 1994; Caselli et al., 1995; Fenson et al., 1994; Gentner, 1982); call this stage the ‘verb spurt period’. (For a review of lexical development and the noun-verb asymmetry, see Gleitman et al., 2005). Given the robust documentation of productive vocabulary development favoring nouns over verbs, it is reasonable to expect that in comprehension, knowledge of the verb-event correspondence may emerge later in development than knowledge of the noun-object correspondence. And in fact, development of the ability to map verbs onto event concepts is documented to be much later: while the noun-object correspondence is formed soon after the first birthday, the verb-event correspondence has not been shown to emerge until late in the second year (Bernal, Lidz, Millotte, & Christophe, 2007; Oshima-Takane, Ariyama, Kobayashi, Katerelos, & Poulin-Dubois, 2011; Waxman et al., 2009).

Waxman et al. (2009) demonstrated that English-learning infants were able to map novel verbs to event categories at 24 months. In a preferential looking task, during familiarization, infants were presented with dynamic scenes (e.g. a man waving a balloon) and some sentences describing the scenes - half of the infants heard descriptive sentences involving a novel verb (e.g. ‘the man is *larping* a balloon’), and the other half heard sentences involving a novel noun (e.g. ‘the man is waving a *larp*’); at test, they were shown two different scenes on opposite sides of the screen – a familiar scene (e.g. a man waving a balloon) and a different scene with a change in the action (e.g. a man tapping a balloon), and were asked to look at the screen by a *baseline* prompt (i.e. ‘now, look at them’) and then by a *response* prompt – those who heard a novel verb in familiarization were prompted with a question about the novel verb (e.g. 'where's *larp*ing?'), and those who heard a novel noun in familiarization were prompted with a question about the novel noun (e.g. 'where's the *larp*?'). Results were: all infants looked more to the novel scene during the baseline window, but only those infants in the novel verb condition later overcome this novel preference and shifted their look to the familiar scene during the response window, whereas those in the noun condition maintained the novelty preference after hearing the response prompt. These results suggested that these 24-month-olds were able to distinguish verbs and nouns, had knowledge about the conceptual categories verbs and nouns corresponded to - in particular, they considered verbs to pick out event categories, not nouns, and were able to use this knowledge to learn novel verb meanings.

Bernal et al. (2007) showed that French-learning infants were able to use the verb-event correspondence in learning novel verb meanings at 23 months. In a preferential pointing task, during familiarization, infants were presented with scenes of an animated object performing self-propelled actions (e.g. a flower rotating) and they heard some sentences describing the scene - half of them heard sentences involving a novel verb (e.g. 'Elle *poune*', meaning "it's *poune*-ing"[11]), and the other half heard sentences involving a novel noun (e.g. 'une *poune*', meaning "it's a *poune*"); at test, they were shown two different scenes on opposite sides of the screen – a familiar scene (e.g. a flower rotating) and a different scene with a change in the action (e.g. a flower jumping), and were asked to point to either of them by a prompt - those in the novel verb condition were asked about

[11] Meanings of sentences are put in double quotation marks, whereas the sentences per se are put in single quotation marks.

the novel verb (e.g. 'laquelle *poune*?', meaning "which *pounes*?"), and those in the novel noun condition were aked about the novel noun (e.g. 'laquelle est une *poune*?', meaning "which is a *poune*?"). Results were: infants in the novel verb condition pointed more towards the familiar scene whereas those in the novel noun condition pointed more towards the different scene. These results suggested that these 23-month-olds knew the verb-event correspondence and exploited it to learn novel verb meanings.

In a more recent work, Oshima-Takane et al. (2011) demonstrated deployment of the verb-event correspondence in learning novel verbs in 20-month-old Japanese-learning infants. In a habituation-switch task, during habituation, infants were presented with two different scenarios many times until they reached some pre-set habituation criterion; in each of the scenes, an object was engaging in some action (e.g. Scene A – an animal bouncing; Scene B – a vehicle jumping); each scene was paired with a sentence containing a novel intransitive verb (e.g. Scene A paired with Sentence A – '*moke*-shi-te(i)ru-yo', Scene B paired with Sentence B – '*seta*-shi-te(i)ru-yo'; see (2-1) for linguistic gloss).

(2-1) *moke/seta*-shi-te(i)ru-yo

moke/seta-do-Present progressive-final particle

"(It) is *moke/seta*-ing."

When infants reached habituation, test began, which for each infant included three switch trials where the original pairings between sentences and scenes were broken: in the word-switch trial, a sentence was paired with the other scene where both the original action and object were switched (e.g. Sentence B, originally paired with vehicle jumping, was switch to pair with animal bouncing (Scene A)); in the action-switch trial, an original sentence was paired with a new scene where the original object was performing a different action (e.g. Sentence B, originally paired with vehicle jumping, was switched to pair with vehicle bouncing); in the agent-switch trial, an original sentence was paired with a new scene where a different object was performing the original action (e.g. Sentence B, originally paired with vehicle jumping, was switched to pair with animal jumping). All infants received the word-switch trial first, followed by the other two switch trials, the order between which were counterbalanced. Infants' performance in the word-switch trial was used to decide if they successfully learned the original pairings presented during habituation: successful dishabituation (i.e. recovery of attention) at the word-switch trial was taken as evidence that an infant learned the original pairing, and thus was categorized as a *recoverer*, otherwise a *non-recoverer*. Results were: recoverers dishabituated to the action-switch trial but not to the agent switch trial, and non-recovered did not dishabituate to either. Because only in the action-switch trial (but not in the agent-switch trial) the novel verb was mapped onto a different event, these results suggested that as long as infants learned the original scene-sentence pairings (i.e. recoverers), they were able to demonstrate their knowledge that verbs mapped to events but not objects. This established the so-far earliest age (i.e. 20 months) when the verb-event bias is seen deployed in learning novel verb meanings.

2.1.2 Motivation for Current Study

The above findings all point to 20-24 months of age as the time when the verb-event correspondence is first seen used to learn novel verbs. This period is closely time-locked with the period when infants' productive vocabulary[12] starts to include an appreciable number of verbs (Fenson et al., 1994; Gentner, 1982; Gleitman et al., 2005; inter alia). Therefore, regarding the relation between knowledge of the verb-event correspondence and growth of verb vocabulary, previous findings cannot tease apart the following possibilities: it could be that the verb-event bias greatly facilitates verb learning thus resulting in the verb spurt, but it could also be that the growth of verb vocabulary supports induction of the verb-event correspondence. In the domain of noun learning, evidence for knowledge the noun-object correspondence is seen by 14 months (Waxman & Booth, 2001), prior to a substantial increase in productive noun vocabulary at 17-18 months (see Gleitman et al., 2005). This makes it not very likely that the noun-object correspondence is generalized from experience because an inductive basis for making generalization may have not yet been built; rather, the link is more likely to have an origin from UG – the noun-object link is specified within the

[12] To be fair, perhaps productive vocabulary is not the best benchmark, but receptive vocabulary is; but since productive vocabulary is the most widely used measure for vocabulary development so far in the field, and receptive vocabulary is not very easy to measure in early infancy, we still rely on the productive measure, while keeping the alternative in mind. See Bergelson and Swingley (2012) for evidence of word comprehension as early as 6 months.

inference engine, and any word that is categorized as a *noun* would trigger the inference chain, leading to an object-kind concept as the target meaning of the word. Similarly, for verb learning, only evidence that the learner exploits the verb-event correspondence *prior to* the verb spurt period would provide strong support for a non-inductive origin of this bias. This motivated the current study to look at a stage before 20 months of age, the earliest documented period of deployment of this bias, as well as the period when verb vocabulary started to grow.

How early should the investigation go? This lower boundary of our investigation window depends on when the prerequisites for this link are established, that is, the linguistic underpinnings that support categorization of *verb* and the conceptual underpinnings that support categorization of *event*.

For the linguistic matters, I have discussed in Chapter 1 that the linguistic underpinnings for verb categorization are largely in place by the first birthday (Mintz, 2006; Peterson-Hicks, 2006) (see Section 1.3.1). Notice, though, 'verb categorization' is a kind of shorthand in talking about this issue, while the most accurate saying should be 'categorization of those words that we know as verbs' are largely in place by the first birthday. For convenience of discussion, I will continue using this shorthand, but it is worth spelling out this important nuance here. What these findings suggest is that the learner knows a) there is this grammatical category G, words belonging to which occur in certain distributions (e.g. 'can __' or 'will __' in English); and b) this category G is separate from another grammatical category G', words of which occur in different distributions (e.g. 'her __' or 'the __' in English); what they do not suggest, but are often taken to suggest, is whether the learner also knows that category G corresponds to what we know as the *verb* category and category G' corresponds to what we know as the *noun* category. The labels do not matter for inductive generalizations: all an inductive learner needs is the observation that words of grammatical category G tend to pick out concepts of category C and that words of grammatical category G' tend to pick out concepts of category C', on the basis of which the learner could discover the G-C and G'-C' links, and subsequently utilizes these links to guide further verb learning. The labels *do* matter, however, for links that are specified within UG: suppose it is *verb–event* and *noun-object* that are terms used in UG; for these links to be corrected utilized to guide inference, in addition to recognizing category G and category G', the learner must also recognize that category G corresponds to the category labeled *verb* by UG, and that category G' corresponds to the category labeled *noun* by UG; if the labels were mixed up, inferences would be misled. Despite this difference, what is common for an inductive learner and for an UG-driven learner is that a category must be formed – a group of words that occur in certain distributional environments (e.g. 'can __', 'will __'); and such a category is shown to be established by the end of the first year. Therefore, for our purpose, we can take the end of first year as a lower benchmark for when the linguistic prerequisites are in place. (For an UG-learner, though, it may take some more time for her to discover that the category is labeled *verb* in UG-vocabulary, because the already-specified *verb-event* link to be activated.)

For the conceptual matters, research on event concept development has provided evidence that by the end of the first year, infants have developed the conceptual foundations to support a range of verb meanings, including verbs encoding moving trajectories (e.g. *fall*), causal results (e.g. *open*), intentional actions (e.g. *get*), transactions (e.g. *give*), as well as psychological states (e.g. *see*) (Buresh, Wilson-Brune, & Woodward, 2006). Infants demonstrate sensitivity to fundamental event relations such as agency and goal-directedness very early in development: as early as 3 months of age, infants are able to detect an actor's goal (Sommerville, Woodward, & Needham, 2005); 5-month-olds attribute goals to human agents as well as to non-human agents with animacy features like self–propulsion (Luo & Baillargeon, 2005); and by 12 months of age, infants are able to interpret and draw inferences about goal-directed behaviors of rational agents (Gergely & Csibra, 2003; Gergely, Nádasdy, Csibra, & Bíró, 1995), and are able to predict the ending of a rational goal-directed motion event based on its beginning (Wagner & Carey, 2005). In addition, there is evidence that young infants are aware of certain types of events. For example, knowledge of containment events is seen as early as 2.5 months: Hespos and Baillargeon (2001) found 2.5- and 3.5-month-olds recognized an object could be lowered inside a container with an open but not a closed top. Infants' sensitivity to causal relations in events have also been shown to develop early: Leslie and Keeble (1987) found that reversal of a causal event elicited more interest than the reversal of a non-causal event in 7-month-olds; Casasola and Cohen (2000) found 14-month-olds were able to distinguish pushing and pulling events that only differ in the causal relations among participants. Although it is hard to know whether there is a super category of event, but we do know from a rich literature on the learner's conceptual development that relations to events and subtypes of events are in place early. Therefore, it stands to reason to assume that the event category is in place early as well.

Taken together, findings on the linguistic and conceptual underpinnings that support categorization of *verb* and *event*, suggest that these prerequisites for utilizing the verb-event link are largely in place around infants' first birthday. This provides the lower boundary of the window of the current investigation, that is, 12 months.

Therefore, the current investigation zooms in on the window between 12 and 20 months; specifically, we look at 14- and 18-month-old English-learning infants.

2.2 Experiment Overlook

In two experiments using the Habituation-Switch Paradigm (Casasola & Cohen, 2000; Werker, Cohen, Lloyd, Casasola, & Stager, 1998; Younger & Cohen, 1986), we examined 14- and 18-month-old English-learning infants' ability to extract morphosyntactic information to categorize novel verbs and identify the event concepts these verbs pick out. In Experiment 1, we habituated infants with two events of an animated animal performing self-propelled actions (a penguin spinning event, and a penguin cartwheeling event), each labeled with a different novel verb embedded in intransitive 'is __ing' frames ('it's *dok*ing', and 'it's *pratch*ing'), and tested them on a new event-verb combination (e.g. spinning labeled with *doke* at habituation was labeled with *pratch* at test). In this task, successful acquisition of the novel verbs' meanings would require a) ability to use the morphosyntactic information to categorize the novel words as verbs, and b) knowledge that verbs pick out event categories. These would lead infants to form the hypothesis that the verb *doke* picks out the spinning event and *pratch* the cartwheeling event, and consequently, they would be surprised to see a familiar event labeled with a different verb, reflected in a recovery of attention at test. However, strategies other than having the verb-event knowledge may also lead to a recovery of attention at test; Experiment 2 addressed one such alternative. To preview, results from the two experiments together provided evidence that the verb-event bias was in place at least by 18 months of age, a couple of months prior to the verb spurt period, thus lending support for a non-inductive origin of this bias; at the same time, no evidence from this study was conclusive about infants' knowledge of the verb-event correspondence at 14 months, calling for future studies.

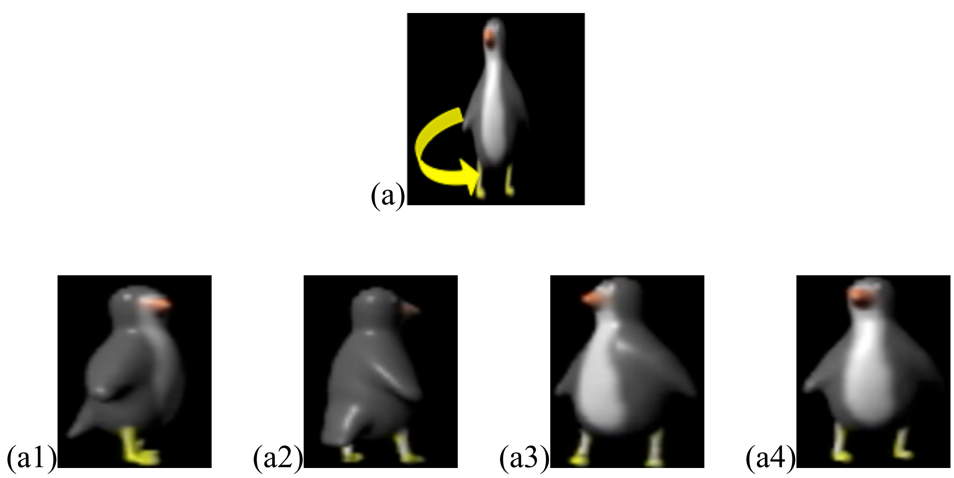
2.3 Experiment 1

2.3.1 Participants

Fourty-two English-speaking infants (21 boys, 21 girls) with a mean age of 14;2 (range: 13;19-14;16) and thirty-four English-speaking infants (17 boys, 17 girls) with a mean age of 18;00 months (range: 17;15-18;16) participated in this experiment. Eighteen additional infants were tested but excluded from the final sample because of experimental error (1), being unable to finish the experiment (7), and failure to habituate (10). All infants were recruited through the Infant Studies Consortium Database at University of Maryland College Park.

2.3.2 Stimuli

The visual stimuli were computer-animated objects engaged in different self-propelled actions. For example, a penguin engaged in a spinning action (Figure 2.1(a)) and a cartwheeling action (Figure 2.1(b)). Each event lasted 15 seconds and was repeated up to two times per trial, giving a maximum trial length of approximately 30 seconds.



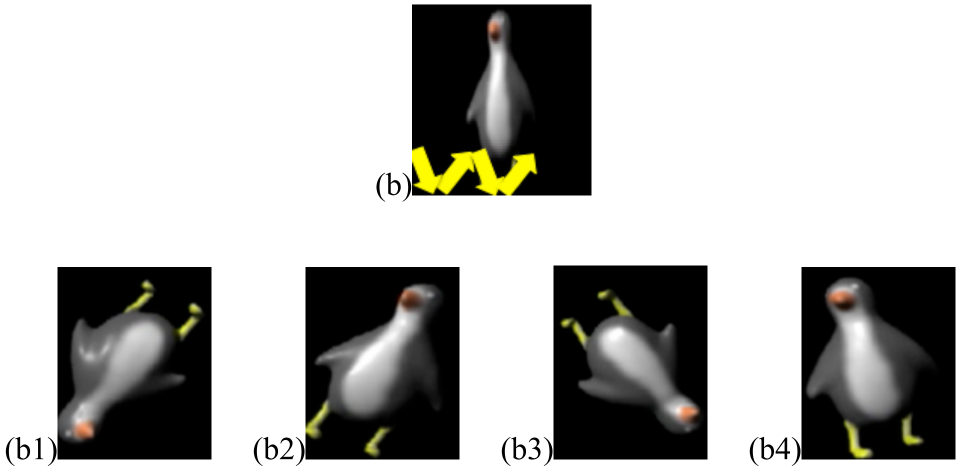


Figure 2.1: Visual stimuli of Experiment 1 & 2 – spinning action (a) & cartwheeling action (b); 1-4 are action snapshots

These visual stimuli were paired with sentences containing novel verbs. Four considerations in selecting the linguistic stimuli were as follows. First, we chose to use intransitive verbs rather than transitive verbs, because they are simpler in argument structure than transitive ones, and the nonlinguistic stimuli required by intransitive verbs are also simpler than those required by transitive verbs; therefore, using intransitive verbs would reduce the potential burden on processing both the linguistic and nonlinguistic stimuli, giving infants a better chance to demonstrate their knowledge of the verb-event correspondence (if any). Second, we chose to use the pronoun *it* to label the animal, also driven by a concern to minimize possible processing burden on infants - some studies (cf. Lidz, Bunger, Leddon, Baier, & Waxman, 2009) reported more processing burden associated with a lexical noun phrase than with a pronoun. Third, the sentences were all in present progressive tense/aspect, because they were presented concurrent with the scenes. Therefore, the linguistic stimuli used were ‘it’s *dok*ing’ and ‘it’s *pratch*ing’; during the 15-second event, the linguistic stimuli were played six times, in slightly different frames: ‘Look, it’s *dok*ing’, ‘Wow, it’s *dok*ing’, ‘Yay, it’s *dok*ing’, ‘Do you see it *dok*ing’.

Stimuli used in this experiment also included an attention-getter, and pre-test and post-test stimuli. The attention-getter stimulus was a video of a butterfly perched on a leaf, which was played if an infant looked away from the screen for more than 1 second, and presented on the screen until the infant’s attention was recaptured. This is a standard adopted by most other studies using this paradigm (e.g. Casasola & Cohen, 2000). Another pair of visual and linguistic stimuli was used at the beginning (pre-test) and end of the experiment (post-test), in which a flower bouncing event was paired with the sentence ‘it’s *sneb*bing’. The duration of this event was the same as the other events (i.e. 15 seconds), and was also played up to two times in its trial (i.e. 30 second maximum). Pre-test and post-test were used to control for fatigue; see Section 2.3.4 for more discussion.

2.3.3 Apparatus

The experiment was run in the Habit version 1.0. program (Cohen, Atkinson, & Chaput, 2004). The stimuli were played on a Samsung wall-mounted 51-inch plasma television, with built-in speakers, located 66 inches away from the chair (or highchair) where the infants were seated. A Sony EVI-D100 video camera was placed directly above the TV monitor. The camera was connected to a color TFT LCD monitor to allow the experimenter to observe the infant’s eye fixation to the screen from a different room, and conduct online coding. Additionally, the video of the child, with a picture-in-picture display of what was on the TV screen, was captured on an iMac computer using QuickTime.

2.3.4 Design

This study used the Habituation-Switch Paradigm (Casasola & Cohen, 2000; Fennell & Werker, 2003; Werker et al., 1998; Younger & Cohen, 1986), the basic design logic of which is this: infants are habituated to two pairs of stimuli (e.g. word A to object A, word B to object B), and are "tested on their ability to detect a switch in the pairing" (Fennell & Werker, 2003) (e.g. word A to object B). With this paradigm, this study tested infants' ability to learn novel verbs from pairs of events and sentences, by first presenting some event-sentence pairs repeatedly over and over again, and when the infant reaches pre-set habituation criterion, a new event-sentence combination was introduced; we measured the amount of attention recovery upon getting the novel combination.

The experiment consisted of the following phases – a pre-test phase, a habituation phase, a test phase, and a post-test phase. The task began with one pre-test trial, where a flower bouncing event paired with 'it's snebbing' was shown. Following that was the habituation phase, during which infants were presented trials of event-sentence pairs for a maximum of 12 times (i.e. a maximum of 12 trials in this phase); they went through the trials until they reached a pre-set *criterion of habituation*, whichever came first. In this experiment, the criterion of habituation, following Werker, Fennell, Corcoran, and Stager (2002), was when an infant's average looking time during any block dropped to less than 65% of average looking time of the most-attended block[13] (i.e. the block that has the longest total looking time); any three consecutive trials made a block. Therefore, the total number of habituation trials each infant received was different. These trials were randomized by blocks of three, to avoid the same event-sentence pair occurring more than two times in a row. Infants who did not meet the criterion of habituation were excluded from the sample of analysis, classified as exclusion due to *failure to habituate*. When infants reached the criterion of habituation or when the 12 trials were all played, whichever came first, the habituation phase was stopped and the test phase began. At test, all infants were presented a fixed number of 2 trials. These two trials were either familiar event-sentence combinations from habituation (*Same condition*), or novel combinations (*Switch condition*). Half of the infants were assigned in the Same condition and the other half in the Switch condition[14]. Each condition had two orders (*Order A* and *Order B*), differing in the order of the two test trials. Following these two test trials, one post-test trial that was the same as the pre-test trial but very distinctive from the habituation and test trials was presented. The

[13] We did not use the first block as the baseline, but used the longest-attended block, for the following reason: we observed that it often took some time (usually one or two trials) for infants to start engaging in the task, so using the mean looking time of the first block may not accurately reflect their baseline attention. Some previous studies used the same method as we did (e.g. Fennell & Werker, 2003; Werker et al., 2002), while some others used the first block (e.g. Oshima-Takane et al., 2011)

[14] We took the dependent variable *condition* as a between-subject variable, rather than within-subject, unlike (Oshima-Takane et al., 2011), for this reason: in a within-subject design, when the Switch trial follows the Same trial, the Same trial is still same as habituation and the Switch is still a switch compared to preceding trials; but when the Switch trial precedes the Same trial, even though Switch still has a switch, Same is no longer strictly speaking same, but could also be considered a switch, relative to the preceding Switch trial.

purpose of having the pre-test and post-test trials was to control for fatigue: if infants were still involved in the experiment towards the end (habituation but not fatigue), we would expect their attention to recover upon seeing the post-test, which was perceptually very distinct from the habituation and test trials. I will discuss how we analyzed habituation controlled for fatigue in Section 2.3.6. See Table 2.1 for a summary of the design.

Stimuli	**Video**	**Audio**
0	flower jumping	*it's snebbing*
1	penguin cartwheeling	*it's pratching*
2	penguin spinning	*it's doking*
3	penguin cartwheeling	*it's doking*
4	penguin spinning	*it's pratching*

(a)

	Same A	**Same B**	**Switch A**	**Switch B**
Pre-test	0			
Habituation	1 & 2 randomized by block of three trials			
Test trial #1	1	2	3	4
Test trial #2	2	1	4	3
Post-test	0			

(b)

Table 2.1: Summary of stimuli in Experiment 1 – (a) visual and auditory stimuli; (b) Stimuli in different condition-order assignments

2.3.5 Procedure and Coding

The procedure began with obtaining the parent(s)' informed consent and collecting the MacArthur Communicative Development Inventory (MCDI) (Dale & Fenson, 1996) – a standardized measurement of productive vocabulary development. When the infant was ready, he/she was led to the test room where the TV monitor and the digital camera were located. The parent came to the test room with the infant and stayed with him/her during the entire process. The infant sat either in the parent's lap or in a highchair in front of the monitor. We took precautions to ensure that the parent could not influence the child's behavior, by explicitly instructing the parent not to direct the infant's attention in any way, and asking the parent to wear a visor (to block sight) in cases where she chose to hold the infants on her lap.

The experimenter began the experiment in the control room next door, by setting up the computer to display an attention-getter (the butterfly). Once the infant looked at the attention-getter, the experimenter pressed the space bar on the computer to begin the first trial, so that the attention-getter on the screen was replaced by the pre-test trial. For each trial, the experimenter pressed a key on the computer when the infant attended to the screen, held the key for as long as the attention was maintained, and released it as soon as the infant looked away. A minimum of 2-second attention was required for it to be counted as a look. A trial continued until the infant looked away for more than 2 continuous seconds or until the end of the trial (approximately 30 seconds). The attention-getter came back on the screen to recapture the infant's attention at the end of each trial and stayed on the screen until the start of next trial. The experimenter was not able to hear the audio, thus unaware of what phase of the experiment the child was in.

Inter-coder reliability was obtained by comparing online coding (as described above) and offline-coding from a second coder on 10% randomly selected infants' recordings from the sample. The Pearson-product moment correlations of the online and offline codings ranged from 0.986 to 0.999, with a mean of 0.997.

2.3.6 Measurement

The dependent variable for analysis was *looking time*, i.e. the amount of time spent on looking at the visual stimuli during a selected window. We used looking time to test for two things: a) successful *habituation* (controlled for fatigue), and b) *dishabituation* at test. To test for habituation, we compared average looking time of the first and last habituation blocks to see if there was a decrease in attention over the habituation trials. To control for fatigue, if infants' attention recovered upon seeing the post-test trial - indexed by a significant increase in average looking time from the last habituation block to post-test, we took that to mean habituation without fatigue. We also compared pre-test and post-test, but did not use this comparison to determine fatigue, because a consistent relation between the two was not established by previous studies – some studies showed that attention in post-test recovered to the same level as in pre-test (e.g. Fennell & Werker, 2003), whereas some showed that post-test attention was no less than

25% of pre-test attention (e.g. Oshima-Takane et al., 2011). To test for dishabituation, we compared average looking time of the last two habituation trials[15] with that of the two test trials; and we took significant increase of looking time from habituation to test as indicator of dishabituation.

2.3.7 Predictions

Infants in the *Same condition* were predicted to show no dishabituation at test. For those in the *Switch condition*, if they learned the associative links between the events and the novel verbs, we expect dishabituation to a familiar event labeled by a different verb (i.e. a novel combination between the event and the verb), reflected by a significant increase in looking time from last habituation trials to test trials. But, if they did not learn the link between the novel word and the event, they were expected to remain habituated. Therefore, such an asymmetry in dependent variable between conditions would suggest infants' ability to learn the verb-event link.

[15] Some studies used mean of the last four habituation trials as the baseline for comparison (Oshima-Takane et al., 2011); and others, had an additional trial following habituation, which played the same event-verb combination, and used looking time of that trial as baseline (e.g. Oshima-Takane et al., 2011).

2.3.8 Results

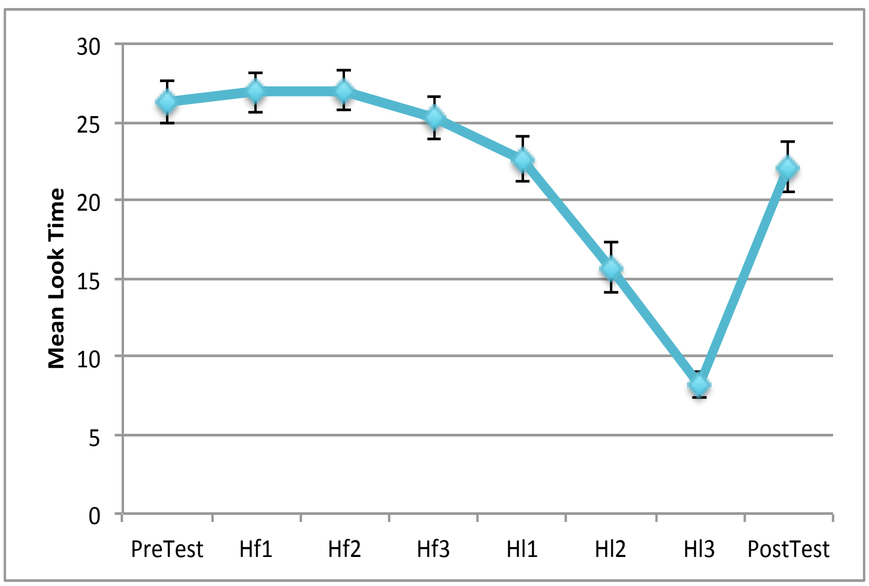
2.3.8.1 Habituation Controlled for Fatigue

Habituation data for the two age groups of participants were analyzed separately. For the 14-month-old group, to determine whether infants were successfully habituated, we conducted planned comparison between the first and last habituation block using one-tailed t-test, and found the mean looking time of the last block ($M = 12.02s$, $SD = 5.00s$) was significantly less than that of the first block ($M = 21.45s$, $SD = 7.55s$), $t(41) = 11.43$, $p < 0.001$. Thus, there was a significant drop in attention throughout the habituation phase. To make sure this habituation was not due to fatigue, we compared post-test to the last habituation block, as planned (one-tailed), and found that the mean looking time of the post-test ($M = 16.77s$, $SD = 10.35s$) was significantly greater than that of the last block ($M = 12.02s$, $SD = 5.00s$), $t(41) = 2.89$, $p = 0.003$. This showed that infants' attention recovered from habituation upon seeing the perceptually distinct post-test trial, suggesting they were not fatigued. The same analyses were conducted for the 18-month-old group, and confirmed that this group was also successfully habituated without becoming fatigued: first, the mean looking time of the last habituation block ($M = 15.51s$, $SD = 3.66s$) was significantly less than that of the first block ($M = 26.4s$, $SD = 15.35s$), $t(33) = 19.99$, $p < 0.001$; second, post-test attention ($M = 22.15s$, $SD = 19.47s$) was significantly greater than that of the last habituation block ($M = 15.51s$, $SD = 3.66s$), $t(33) = 4.43$, $p < 0.001$.

We also compared post-test to pre-test, and found that for both age groups, attention in post-test did not recover to the level of pre-test. Instead, infants generally lost some attention in post-test: 14-month-olds lost about 25% attention in post-test (M = 16.77s, SD = 10.35s) compared to pre-test (M = 21.97s, SD = 9.80s), $t(41) = 2.63$, $p < 0.01$, one-tailed; and 18-month-olds lost about 15% attention in post-test (M = 22.15s, SD = 0.84s) compared to pre-test (M = 26.31s, SD = 9.47s), $t(33) = 2.15$, $p = 0.02$, one-tailed. In previous studies, Fennell and Werker (2003) found that their 14-month-old participants' post-test attention recovered to pre-test level, whereas Oshima-Takane et al. (2011) found that their 20-month-old participants lost about 25% attention in post-test compared to pre-test. Given the inconsistent findings on pre-test and post-test relation, and also because the presence of pre-test might have made post-test less novel and thus less interesting, we decided to remove the pre-test phase in Experiment 2.

The above results were illustrated in Figure 2.2(a) and 2.2(b). Clearly, infants in both age groups began the task with a relatively high level of attention (i.e. pre-test); during the habituation phase, there was a clear decreasing trend from the beginning (Hf1, Hf2 and Hf3 stand for first 3 habituation trials, respectively) to the end (Hl1, Hl2 and Hl3 stand for last 3 habituation trials, respectively, with Hl3 being the very last one); and attention recovered to some extent in post-test, but not back to the same level as in pre-test.

(a)



(b)

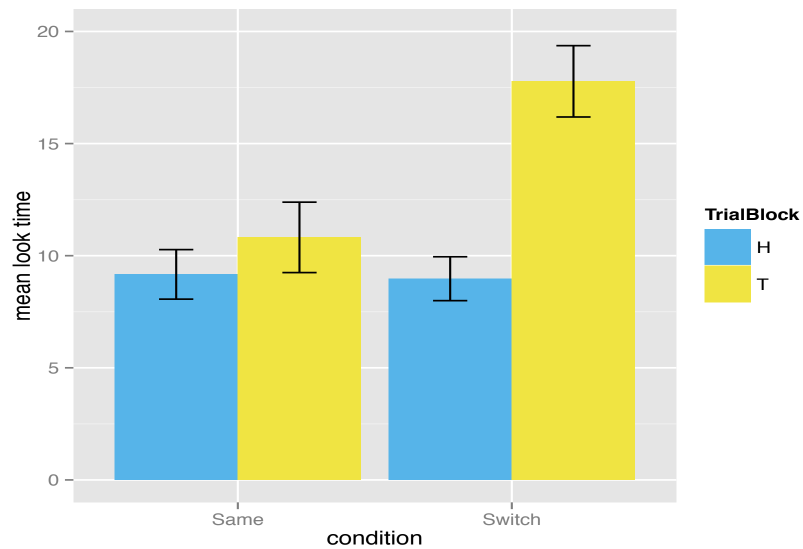
Figure 2.2: Habituation timelines of 14-month-olds (a) **& 18-month-olds** (b)

2.3.8.2 Dishabituation Analysis

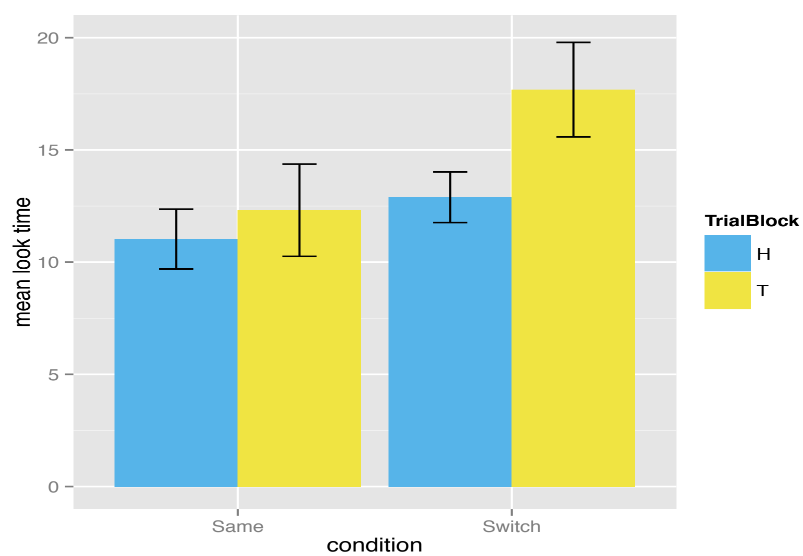
Having determined that infants were successfully habituated, we then conducted the main set of analyses on their performance at test. Data were first entered into a three-way mixed ANOVA with *age group* (14- vs. 18-month-olds) and *condition* (Same vs. Switch) as between-participants factors, and with *trial block* (last two habituation trials vs. two test trials) as a within-participants factor. This analysis revealed a main effect of condition, $F(1, 72) = 9.64$, $p < 0.01$; a main effect of trial block, $F(1, 72) = 18.00$, $p < 0.001$; and a significant interaction between condition and trial block, $F(1, 72) = 7.61$, $p < 0.01$. Specifically, for *Same condition*, attention during test (M = 11.49s, SD = 7.72s) was not significantly different from that in the last two habituation trials (M = 10.00s, SD = 5.27s), $t(37) = -1.13$, $p = 0.26$, two-tailed; for *Switch condition*, however, attention during test (M = 17.74s, SD = 7.83s) was significantly greater than that in the last two habituation trials (M = 10.73s, SD = 14.91s), $t(37) = -4.64$, $p < 0.001$, one-tailed. There was no main effect of age group, $F(1, 72) = 2.53$, $p = 0.12$; and no other interactions. These results suggested the following points: first, 14- and 18-month-olds did not demonstrate different performance patterns; second, infants in this experiment dishabituated to the test stimuli only in the *Switch condition*, but not in the *Same condition*.

Next, data were analyzed separately for the two age groups, each with a two-way mixed ANOVA with *trial block* (last two habituation trials vs. two test trials) as within-subject factor and *condition* (Same vs. Switch) as between-subject factors. For 14-month-old infants, there was a main effect of condition,

$F(1, 40) = 6.15$, $p = 0.02$; a main effect of trial block, $F(1, 40) = 15.83$, $p < 0.01$; and a significant interaction between condition and trial block, $F(1, 40) = 7.41$, $p < 0.01$. Specifically, for the *Same condition*, attention during test (M = 10.82s, SD = 7.29s) was not significantly different from the last two habituation trials (M = 9.17s, SD = 5.06s): $t(20) = -0.86$, $p = 0.40$, two-tailed; but for the *Switch condition*, attention during test (M = 17.78s, SD = 7.29s) was significantly greater than that of the last two habituation trials (M = 8.97s, SD = 4.49s), $t(20) = -4.88$, $p < 0.001$. For 18-month-old infants, there was a marginally main effect of condition, $F(1, 32) = 3.77$, $p = 0.06$; and a marginally main effect of trial block, $F(1, 32) = 3.89$, $p = 0.06$. There was no interaction between condition and trial block, $F(1, 32) = 1.30$, $p = 0.26$. A closer look into each condition revealed the following: for the *Same condition*, attention during test (M = 12.31s, SD = 8.47s) was not significantly different from that during the last two habituation trials (M = 11.03s, SD = 5.49s), $t(16) = -0.7112$, $p = 0.49$, two-tailed; but for the *Switch condition*, attention during test (M = 17.69s, SD = 8.69s) was significantly greater than that during the last two habituation trials (M = 12.89s, SD = 4.64s), $t(16) = -1.92$, $p = 0.04$, one-tailed. See Figure 2.3(a) and 2.3(b) for illustrations of the results.



(a)



(b)

Figure 2.3: Mean look time across trial blocks in different conditions for 14-month-olds (a) **& 18-month-olds** (b)

2.3.9 Discussion

The results of Experiment 1 showed that both 14-month-olds and 18-month-olds dishabituated in the *Switch condition* but not in the *Same condition*. This asymmetry between conditions was consistent with the hypothesis that they were able to recruit morphosyntactic cues online to categorize novel verbs and map novel verbs to event concepts. However, the results are also consistent with an alternative explanation: infants were using some strategy that did not require syntactic analysis or categorization to succeed in the task. Such a strategy would allow infants to track mismatches of any kind, triggering recovery of attention in the *Switch condition*. For example, if infants represented the audio as a whole (as opposed to analyzing it into syntactic units), and represented the video as a whole (as opposed to analyzing it into an event and a participant), and tracked the connection between them – i.e. establishing a audio-video link, instead of a verb-event link; once the original connection was broken (e.g. audio A was originally linked to video A, but now linked to video B), there was a recovery of attention. Therefore, this strategy might have led to the same overall pattern observed in Experiment 1. Experiment 2 was designed to address this possibility.

To disentangle the general-purpose mismatch-detecting strategy from true linguistic categorization based on the verb-event link, there needs to be a condition where the audio-video connection gets switched, but the word-concept mapping remains the same. If infants show no recovery of attention in this condition, then this would suggest they were not using mismatch-detecting strategy in Experiment 1. In Experiment 2, we added such a condition.

2.4 Experiment 2

2.4.1 Participants

Forty-three English-speaking infants (21 boys, 22 girls) with a mean age of 14;5 months (range: 13;12 - 14;20) and Forty-one English-speaking infants (20 boys, 21 girls) with a mean age of 18;1 months (range: 17;12 - 18;19) participated in this experiment. Twenty-five additional infants were tested but excluded from the final sample because of experimental error (3), being unable to finish the experiment (6), parental interference (1), and failure to habituate (15). All infants were recruited through the Infant Lab Database at University of Maryland College Park.

2.4.2 Stimuli

The visual stimuli were same as those used in Experiment 1 - a penguin engaged in a spinning action (Figure 2.1(a)) and a cartwheeling action (Figure 2.1(b)). The linguistic stimuli paired with these visual stimuli were simple intransitive sentences, varying with conditions. For example, there were sentences containing an intransitive novel verb labeling the action (e.g. 'it's *dok*ing'), and sentences containing a novel noun labeling the animal performing the action (e.g. 'it's a *pratch*'). The pairing between the visual stimuli and linguistic stimuli were counterbalanced such that one group of infants heard the verb *doke* label the spinning event and the noun *pratch* labeling the penguin, while the other group heard the verb *pratch* label the spinning event and the noun *doke* labeling the penguin. The same attention-getter stimulus and post-test stimuli were used as those in Experiment 1 (pre-test was removed).

2.4.3 Apparatus

The experimental set-up and apparatus were same as those used in Experiment 1.

2.4.4 Design

This experiment also used the Habituation-Switch Paradigm (Casasola & Cohen, 2000; Werker et al., 1998; Younger & Cohen, 1986), adopting a similar design logic. A major change in Experiment 2 was the introduction of a third condition such that there were three conditions at test - the *Same condition*, where a familiar audio-video combination from habituation was presented; the *Noun-Switch condition*, where the audio with a novel noun was paired with a different event but the same object; and the *Verb-Switch condition*, where the audio with a novel verb was paired with a different event but the same object. These two types of switch conditions were designed to tease apart the general-purpose mismatch-detecting strategy from true linguistic knowledge, because both conditions involve a switch in the pairing between the video and the audio, but only the *Verb-Switch condition* involves a change in word meaning (since the novel verb is used to label a different event). Infants were randomly assigned to one of the three conditions. See Table 2.2 for an example illustrating the design in Experiment 2.

Stimuli Number	Video	Audio
0	flower jumping	it's snebbing
1	penguin cartwheeling	it's pratching
2	penguin spinning	It's a doke
3	penguin cartwheeling	It's a pratch
4	penguin spinning	It's doking
5	penguin cartwheeling	It's a doke
6	penguine spinning	it's pratching
7	penguin cartwheeling	It's doking
8	penguin spinning	It's a pratch

(a)

Condition Order	Same A	Same B	Same C	Same D
Habituation	1 & 2	1 & 2	3 & 4	3 & 4
Test trial #1	1	2	3	4
Test trial #2	1	2	3	4
Post-test	0	0	0	0

(b)

Condition Order	Noun-Switch A	Noun-Switch B	Verb-Switch A	Verb-Switch B
Habituation	1 & 2	3 & 4	1 & 2	3 & 4
Test trial #1	5	8	6	7
Test trial #2	5	8	6	7
Post-test	0	0	0	0

(b Cont.)

Table 2.2: Summary of stimuli in Experiment 2 – (a) visual and auditory stimuli; (b) stimuli in different condition-order assignments

On top of this major difference in design, Experiment 2 also differed from Experiment 1 in a few other aspects: First, Experiment 2 did not have a pre-test phase, for two reasons: a) we found in Experiment 1 that infants' attention during post-test did not recover to the same level as pre-test, and previous studies did not establish a consistent ratio (between post-test and pre-test looking time) to be used as a threshold to determine fatigue; and more importantly, b) the presence of pre-test stimuli might make the post-test stimuli less novel to infants, making it hard to say if the decreased attention in post-test was due to fatigue or less interest in a familiar stimuli. Second, the two test trials in Experiment 1 were different, whereas those in Experiment 2 were the same trial presented twice; this was because the two ways of switching in Experiment 2 were different in nature – one maintained the meaning of the novel word (*Noun-Switch*), the other changed the meaning (*Verb-Switch*).

2.4.5 Procedure and Coding

The procedure was the same as Experiment 1, except for absence of pre-test phase. Inter-coder reliability was obtained by comparing online coding (as described above) and offline-coding from a second coder on 10% randomly selected infants' recordings from the sample. The Pearson-product moment correlations of the online and offline codings ranged from 0.993 to 0.999, with a mean of 0.997.

2.4.6 Measurement

As in Experiment 1, we used *looking time* as dependent variable to test for habituation and dishabituation.

2.4.7 Predictions

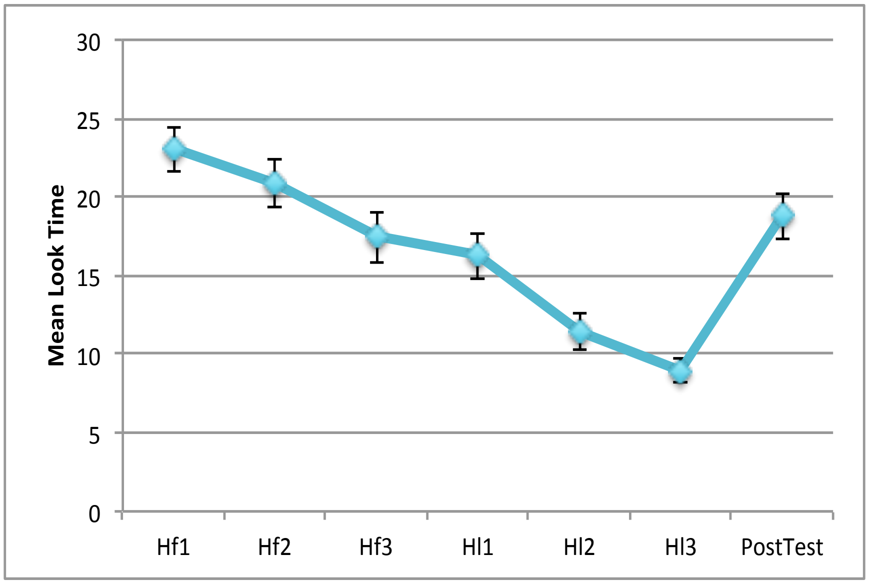
The main purpose of Experiment 2 was to examine the possibility of infants using a general-purpose mismatch-detecting strategy to dishabituate to any type of switch, regardless of meaning. If this were the case, then we would expect to see dishabituation in both *Noun-Switch* and *Verb-Switch* conditions – that is, a significant increase of looking time from last two habituation trials to test, and no dishabituation in *Same condition*. If, however, they were not using such a strategy, but were truly analyzing the linguistic stimuli into meaningful units and knew the verb-event and noun-object mappings, then, we would expect to see dishabituation only in the *Verb-Switch condition* where meaning was changed, but no dishabituation in the *Noun-Switch* or the *Same condition*.

2.4.8 Results

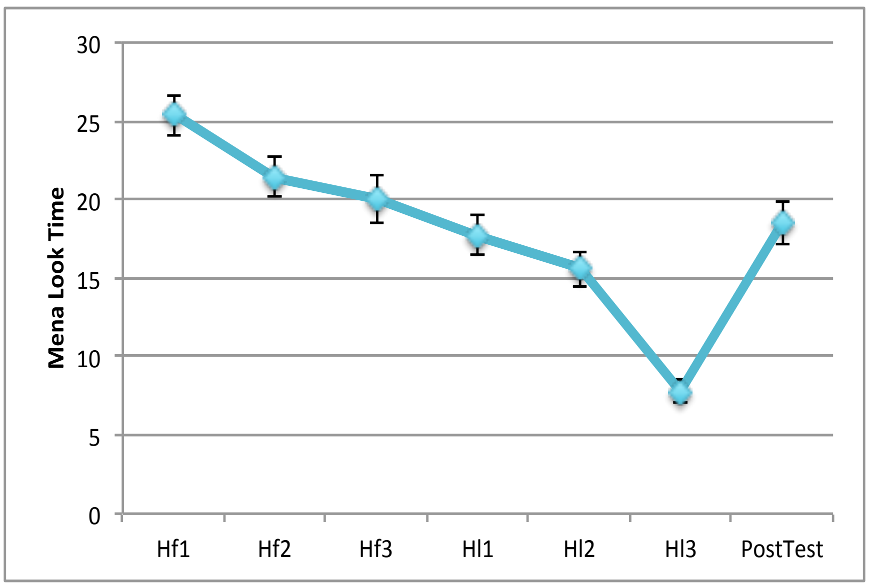
2.4.8.1 Habituation Controlled for Fatigue

As in Experiment 1, to determine whether infants were successfully habituated, we conducted planned comparison between the first and last habituation block using one-tailed t-test. If there was habituation, then we compared post-test to the last habituation block to make sure this habituation was not due to fatigue, also using one-tailed t-test. Habituation data for the two age groups of participants were analyzed separately. For 14-month-olds, the mean looking time of the last block (M = 12.23s, SD = 4.82s) was significantly less that that of the first block (M = 20.46s, SD = 7.19s), $t(42) = 12.41$, $p < 0.001$, indicating a significant drop in attention throughout habituation phase; and the mean looking time of post-test (M = 18.80s, SD = 9.66s) was significantly greater than that of the last block (M = 12.23s, SD = 4.82s), $t(42) = -4.51$, $p < 0.001$, assuring us that infants' habituation was not due to fatigue. For 18-month-olds, the mean looking time of the last habituation block (M = 13.70s, SD = 3.43s) was significantly less than that of the first block (M = 22.30s, SD = 6.19s), $t(40) = 13.20$, $p < 0.001$; and post-test attention (M = 18.49s, SD = 8.59s) was significantly greater than that of the last habituation block (M = 13.70s, SD = 3.43s), t(40) = , $p < 0.001$; indicating habituation without fatigue.

The above results were illustrated in Figure 2.4(a) and 2.4(b). Clearly, infants in both age groups began the task with a relatively high level of attention, which decreased over habituation trials; and attention recovered to some extent in post-test.



(a)



(b)

Figure 2.4: Habituation timelines for 14-month-olds (a) **&18-month-olds** (b)

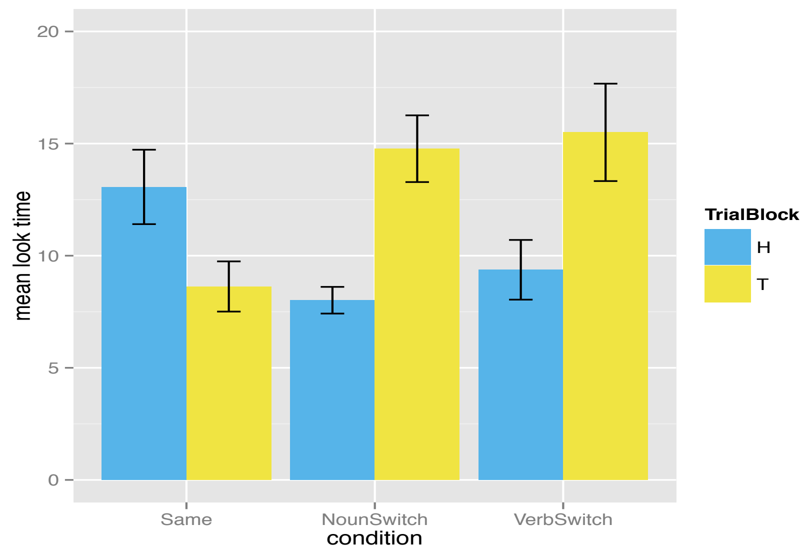
2.4.8.2 Dishabituation analysis

Having determined that infants were successfully habituated, infants' performance at test was analyzed. Data were first entered a three-way mixed ANOVA with *age group* (14- vs. 18-month-olds) and *condition* (Same vs. Noun-Switch vs. Verb-Switch) as between-participant factors, and with *trial block* (last two habituation trials vs. two test trials) as a within-participant factor. This analysis revealed a main effect of condition, $F(1, 78) = 3.94$, $p = 0.02$; a main effect of trial block, $F(1, 78) = 8.46$, $p < 0.01$; a significant interaction between condition and trial block, $F(2, 78) = 24.33$, $p < 0.001$; and a significant interaction between age group, condition, and trial block, $F(2, 78) = 3.61$, $p = 0.03$. This three-way interaction invited us to look closer into the two age groups to see how their performance patterns differed. Therefore, data were then analyzed separately for the two age groups, each with a two-way mixed ANOVA with trial block (last two habituation trials vs. two test trials) as a within-participant factor and condition (Same vs. Noun-Switch vs. Verb-Switch) as a between-participant factor.

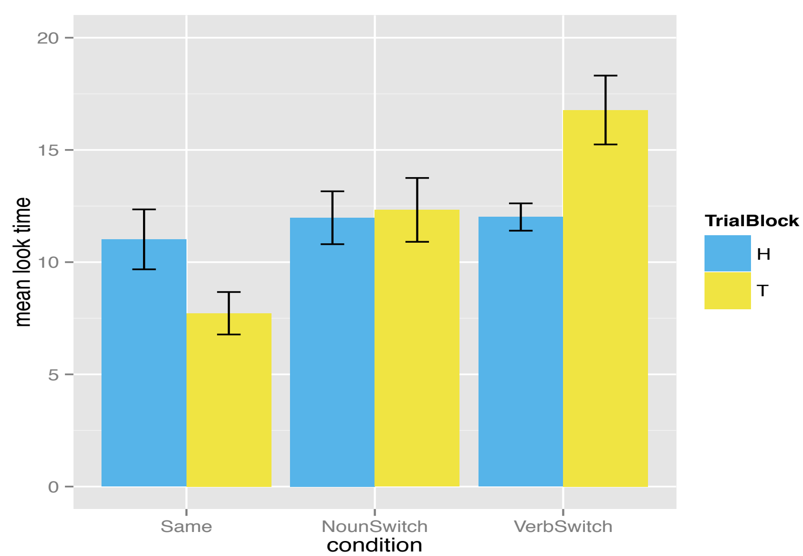
For 14-month-olds, there was a main effect of trial block, $F(1, 40) = 7.50$, $p < 0.01$; and a significant interaction of condition and trial block, $F(2, 40) = 14.43$, $p < 0.001$. Specifically, for the *Same condition*, attention during test ($M = 8.63s$, $SD = 4.33s$) was significantly less than that of the last two habituation trials ($M = 13.07s$, $SD = 6.43s$), $t(14) = 3.00$, $p < 0.01$, two-tailed; for *Noun-Switch condition*, attention during test ($M = 14.77s$, $SD = 5.57s$) was significantly greater than that habituation ($M = 8.01s$, $SD = 2.23s$), $t(13) = -4.35$, $p < 0.001$; and for

Verb-Switch condition, test ($M = 15.5s$, $SD = 8.13s$) has a longer looking time than habituation ($M = 9.37s$, $SD = 4.98s$), $t(13) = -3.11$, $p < 0.01$. These suggested that 14-month-old infants dishabituated to test stimuli in both the *Noun-Switch* and *Verb-Switch conditions*, but not in the *Same condition*. See Figure 2.5(a) for illustration.

For 18-month-olds, there was a main effect of condition, $F(2, 38) = 5.305$, $p < 0.01$; and a significant interaction between condition and trial block, $F(2, 38) = 13.376$, $p < 0.001$. Specifically, for the *Same condition*, attention during test ($M = 7.72s$, $SD = 3.41s$) was significantly less than that during the last two habituation trials ($M = 11.02s$, $SD = 4.81s$), $t(12) = 3.56$, $p < 0.01$, two-tailed; for the *Noun-Switch condition*, attention during test ($M = 12.33s$, $SD = 5.32s$) was not significantly different than that of habituation ($M = 11.98s$, $SD = 4.41s$), $t(13) = -0.33$, $p = 0.37$; and for *Verb-Switch condition*, test ($M = 16.78s$, $SD = 5.74s$) had a significantly longer looking time than habituation ($M = 12.01s$, $SD = 2.28s$), $t(13) = -3.78$, $p = 0.001$. These suggested that 18-month-olds dishabituated to test stimuli only in *Verb-Switch condition*, but not in *Noun-Switch* or *Same condition*. See Figure 2.5(b) for illustration.



(a)



(b)

Figure 2.5: Mean look time across trial blocks in different conditions for 14-month-olds (a) **& 18-month-olds** (b)

2.4.9 Discussion

The results of Experiment 2 showed that 14-month-olds dishabituated to both types of switch, whereas 18-month-olds only dishabituated to the type of switch that had a change in word meaning (i.e. *Verb-Switch*). These findings suggested the two age groups may have used different strategies in completing our switch-tasks. While 18-month-olds may be analyzing the linguistic stimuli into meaningful units, recruiting morphosyntactic cues online to categorize novel words (i.e. use *is __ing* dependency to category verbs), and mapping novel words onto their corresponding event categories (i.e. verb-event, noun-object), their four-month younger peers may have simply attended to the surface-level connection between the video and audio stimuli. For the 14-month-olds, therefore, whenever there was a change in the connection between the video and audio stimuli (i.e. *Noun-Switch* and *Verb-Switch conditions*), their attention was recaptured. Eighteen-month-olds, however, *only* re-attended to the stimuli when there was a change in terms of word meaning (i.e. *Verb-Switch condition*).

2.5 General Discussion

To summarize, the study presented from Section 2.2 to Section 2.4, with two experiments using the Habituation-Switch Paradigm, demonstrated that English-learning 18-month-old infants were able to rapidly learn the meanings of novel verbs from presentations of simple motion events performed by an animated agent, paired with novel intransitive verbs embedded in 'is __ing' frames, but provided no conclusive evidence for 14-month-olds's ability to do so. In what follows, I will discuss the implications of these results (Section 2.5.1), and then point out some caveats in interpreting the results and some directions for future work (Section 2.5.2).

2.5.1 Implications

Word learning is an impressive ability because the context of use makes available infinitely many possible meanings for any given word form, and therefore calls for guided learning. One such guidance proposed come from certain correspondences between form and meaning, for example, the mapping between the grammatical category *verb* and the conceptual category *event*. Having the verb-event correspondence in place early in development can aid learning novel verb meanings by restricting the conceptual categories that provide candidate verb meanings.

We see evidence that infants may have the linguistic and conceptual underpinnings that support establishment of this correspondence by the end of first year (Buresh et al., 2006; Casasola & Cohen, 2000; Gergely & Csibra, 2003; Gergely et al., 1995; Hespos & Baillargeon, 2001; Leslie & Keeble, 1987; Luo & Baillargeon, 2005; Mintz, 2006; Sommerville et al., 2005; Wagner & Carey, 2005). However, previous findings do not show the verb-event correspondence being deployed for novel verb learning until quite late in the second year: 20 months for Japanese-learning infants (Oshima-Takane et al., 2011), 23 months for

French-learning infants (Bernal et al., 2007), and 24 months for English-learning infants (Waxman et al., 2009). This time period is also the same time that we see an appreciable number of verbs start to be included in infants' productive vocabulary (Fenson et al., 1994; Gleitman et al., 2005; inter alia). Given that infants' use of this correspondence to learn novel verbs was tightly correlated in time with growth in verb vocabulary, it is unclear whether the growth of a verb vocabulary supports the induction of the verb-event link or whether the activation of the verb-event link already specified in UG supports rapid growth in verb vocabulary.

The current study pushes down the age at which infants are able to use the verb-event correspondence for novel verb learning to 18 months of age. In turn, this observation implies that deployment of a potential heuristic for verb learning – the verb-event bias, is in place a few months before the substantial increase of verb usage, suggesting that the verb-event correspondence allows for the growth of the verb vocabulary, and not the other way round. Therefore, regarding the question about the origin of the verb-event bias, findings of this study makes an inductive origin less likely, and lends more credibility to an origin from the language acquisition device (or, UG). Specifically, what an UG origin means is this: links between certain grammatical categories and certain conceptual categories may be specified within the inference engine of UG, for example, the verb-event link; once the learner discovers the surface correlates of 'verbhood' in the particular language she is learning - that is, she is able to categorize a novel word into the *verb* category, the inference chain from *verb* to *event* is automatically activated.

When is the inference chain first activated? We do not know. But what we do know from current results is that it certainly is activated by 18 months. As for the other earlier stage tested, 14 months of age, the current study does not have evidence for the verb-event bias being deployed, which might imply one of two possibilities: a) the verb-event link is not yet activated by 14 months – although a category for what we know as verbs is shown established by 12 months, infants may still need more time to figure out the label of that category is *verb* in UG terms (see Section 2.1.2); or b) the verb-event link is already activated by 14 months, but is not demonstrated in current tasks; see more discussion on this in the next section

Another implication of this study, in addition to answering the central question about the origin of the verb-event bias, is that it augments the literature on learning from morphosyntactic cues. Previous studies have shown young English-learning infants are sensitive to morphological cues to verbs (Santelmann & Jusczyk, 1998), and also documented infants' ability to use the cues for novel verb categorization (Mintz, 2006; Peterson-Hicks, 2006). The present study adds to this literature by showing that young infants (at least 18-month-olds) are also able to utilize this morphosyntactic-based categorization in learning novel verb meanings.

2.5.2 Caveats and Future Directions

Three caveats in interpreting the results are worth pointing out.

The first caveat is about 14-month-olds' failure. To successfully learn verb meanings in these tasks, it required infants to have the knowledge, and to be able to deploy their knowledge, in at least two respects: first, they should be aware of the frequent frame 'is __ing' as a context for verbs to occur in, and they should be able to recruit this morphosyntactic information online to categorize novel verbs; and second, they should be aware of the link between the grammatical category *verb* and the conceptual category *event*, and be able to use this general correspondence to quickly learn the specific links presented in the task. Lack of any of the above knowledge/abilities could lead to failure in verb learning in the current study. Therefore, we could not conclude from the results that 14-month-olds do not yet have the knowledge of the verb-event correspondence.

Their performance in this study *could* indicate lack of the verb-event link, but could *also* be a reflection of failure in any of the above-mentioned respects: failure to recognizing the 'is __ing' dependency as a verb frame, failure to use this frame for verb categorization online, or failure to deploy the verb-event link to learn novel verb meanings within the experimental context. All these aspects need to be independently tested. For example, although there is evidence that 12-month-olds and 15-month-olds are able to use morphosyntactic cues like preceding auxiliaries (e.g. 'can __', 'will __') to categorize novel words as what we know as verbs (Mintz, 2006; Peterson-Hicks, 2006), whether the 'is __ing' frame can effectively cue to the same category still lacks direct evidence in general, nor for infants at 14 months in particular. The only evidence related to the 'is __ing' frame is about infants' sensitivity to this frame as a well-formed morphosyntactic dependency – that 18-month-olds could distinguish 'is __ing' from 'can __ing' sentences (Santelmann & Jusczyk, 1998); but this does not suggest infants are able to use this dependency for verb categorization at 18 months, nor at 14 months.

Moreover, 14-month-olds' failure in any of the above-mentioned respects may represent a true lack of knowledge (e.g. not knowing the verb-event correspondence, not knowing *is__ing* is a verb context), but may as well be a failure in the task designed to reveal their knowledge. For example, it is possible that the current design, by repeating the same token events several times rather than showing several tokens of the same type, failed to promote categorization of the events. Future work may examine the role of variability in promoting event categorization in 14-month-olds.

A second caveat is that there may be a leaner interpretation of 18-month-olds' successful verb learning in our task. Suppose that 18-month-olds know the noun-object correspondence and are able to categorize nouns and verbs using distributional cues, but that they do not yet know the verb-event correspondence. In this case, infants could still succeed in the current study by inferring that novel words that are not nouns do not pick out object concepts. Thus, by noticing that *doke* in 'it's *dok*ing' is not a noun, infants infer that it must refer to some other aspect of the scene, in this case the event. This kind of strategy could in principle enable infants to *appear* to know the verb-event correspondence, even without knowledge of it. We think this possibility is not very likely, however. In the adjective learning literature, this pattern of inference has not been observed. In experiments examining the correspondence between adjectives and object properties, knowledge of the noun-object correspondence does not force infants to link an adjective to a property concept when the property concept is sufficiently salient in the scene (Waxman & Booth, 2001). Nonetheless, future work should attempt to test this possibility in the domain of verb learning.

Last but not least, I have been using infants' *productive* vocabulary as the basis to define the verb spurt period – a period when infants start to *produce* lots of verbs, which is in turn used as a benchmark to evaluate the two hypotheses about the origin of the verb-event bias - inductive generalization versus UG specification. My argument, in a nutshell, goes like this: current results suggest infants are able to utilize the verb-event link to learn novel verb meanings by 18 months of age, which is a couple of months *prior to* a significant growth of their verb vocabulary; this means by the time they are able to deploy this link, an inductive basis for generalization is not likely to be in place. My argument, however, is based on the assumption that productive vocabulary is a good indicator of the adequacy of an inductive basis. This is a fair assumption, I believe, given the inaccessibility of large-scale receptive verb vocabulary data, but a more accurate evaluation would require: a) pinpointing the verb spurt period in terms of receptive vocabulary, b) comparing the receptive spurt period to the productive spurt period, c) evaluating whether each type of spurt is adequate for an inductive basis to be built, and d) locating the earliest time of deploying the verb-event link in relation to the time when an adequate inductive basis is believed to be established.

Chapter 3

Specificity/Generality of Expected Verb Meanings

Verbs in general all describe *the way* in which the participants are *involved* in events, but not all verbs also describe what *kinds of things* the participants are. Some verbs place specific semantic constraints on their arguments: for example, one can only *eat* things that are edible and *drink* things that are liquidy; only things with minds can *think*, one horses can *gallop*, and only animals can *hibernate*; ec cetera. These verbs are more specific in meanings, and are more narrowly extendible - only to a restricted set of participants; see (3-1) for an example. Some verbs, on the other hand, place little semantic constraint on their arguments: for example, anything can *fall* or *spin*, anything can be *dropped* or *kicked*, ec cetera. These verbs are more general in meanings, and are more broadly extendible; see (3-2) for an example.

(3-1) [[hibernate]] = λe. λx. hibernate(e) & Agent(e, x) & animal(x)

(3-2) [[spin]] = λe. λx. spin(e) & Agent(e, x)

(3-3) [[gorp]] = λe. λx. V(e) & Agent(e, x)

(3-4) [[gorp]] = λe. λx. V(e) & Agent(e, x) & CAT(x)

The verb-event bias says the learner expects verbs to describe kinds of events. A closely related question is: what kinds of events does the learner expect a novel verb to describe; in other words, how specific/general does the learner expect the event concept of a novel verb to be[16]? Suppose a novel verb *gorp* is heard in the context of a cat involved in an event *e* as an agent, does the learner expect *gorp* to describe the set of *e* events the agent of which has little restriction (as in (3-3)), or the set of *e* events that are specifically performed by cat sorts of things (as in (3-4))? Here what seems a plausible conjecture is: the learner's initial expectation about verb meanings is more general than specific, and finer semantic constraints of particular verbs are later added.

To answer this question – to what extent the learner's initial expectations of verb meanings are specific versus general, a commonly used way of investigation is to examine young learners' verb extension ability. If the learner starts with a general expectation, she would expect verb meanings to be broadly extendible; then, hearing the verb *gorp* used in an instance of a cat involved in an event *e*, the learner is likely to readily accept a horse involved in the same event as an agent to be labeled by *gorp* too. If, on the other hand, the learner's initial expectation is specific, she would expect verb meanings to be narrowly extendible; then, she may infer that *gorp* only applies to cat sorts of things, leading to rejection to the same event with a horse as agent to be labeled by *gorp*.

[16] Notice that specificity/generality of verb meanings is discussed on a continuum, rather than just two discrete extremes.

Previous findings about young children's verb extension abilities are mixed: some reported even preschoolers were *not* willing to extend a novel verb to the same event category with a different (sort of)[17] participant, suggesting rather specific expectations about verb meanings; whereas some showed young infants were willing to do so. Mixed results were obtained from studies that varied in several non-trivial aspects, including type of participants tested (e.g. extension to agent or patient), type of novel verbs used (e.g. transitive or intransitive), type of arguments (e.g. null, pronominal, or lexical argument) and level of cognitive demands the tasks imposed, ec cetera.

In fact, one important concern in modern developmental psychology is to distinguish real absence of some knowledge from non-showing of it in laboratory setting. Therefore, this chapter aims to examine young learners' verb extension abilities under some conditions where cognitive demands are minimized; in other words, efforts are taken to control for some possible barriers that may prevent young learners from demonstrating their real knowledge. In particular, this chapter presents an experiment testing 23-month-old English-learning infants' willingness to extend a novel *intransitive* verb to the same event category with a different sort of *agent*, with a particular manipulation to control for the processing loads imposed by different types of subject NPs (e.g. pronominal vs. lexical). Results show that these infants were willing to extend a newly learned verb to the same event type with a different type of agent, when the novel verb was taught in

[17] In some of these studies, not only children were unwilling to extend to a different *sort of* participant, but they were even unwilling to accept extension to a different participant *of the same sort* (e.g. original agent was person A; person B performing the same action was not accepted as an instance of the same verb.)

a sentence with a pronominal subject. These findings lend more support for the conjecture that infants begin with a rather general expectation about verb meanings.

In what follows, I will begin with a review of mixed findings from previous research on young learners' verb extension abilities, and some findings that motivated the subject type manipulation of the current experiment (Section 3.1); then, I will discuss the hypotheses of the current study and its overall design (Section 3.2); following that I will present an experiment designed to test these hypotheses (Section 3.3); from there, I will discuss what the findings imply about the central question of this chapter – to what extent the learner expects verb meanings to be general/ specific (Section 3.4).

3.1 Background

Section 3.1.1 first reviews mixed previous results with regard to young learners' verb extension ability, which inspire the discussion about the importance to separate lack of knowledge from failure to demonstrate some knowledge due to task demands; and then Section 3.1.2 reviews some results that showed young infants were sensitive to processing load incurred by different types of linguistic stimuli.

3.1.1 Mixed Findings about Children's Verb Extension Abilities

Studies probing young learners' expectations about the specificity/generality of verb meanings usually involve teaching them some novel verb in the context of some event with certain objects/people as participants, and examining their ability to extend this newly learned verb to a different instance of the same event category, varying its participant(s) while keeping the action/motion constant. Success in verb extension is taken as evidence for a larger degree of generality in the learner's early expectation about verb meanings, whereas failure is considered to indicate a larger degree of specificity. Below I review some past findings on children's verb extension ability.

3.1.1.1 Some Cases of Failure in Verb Extension

The findings usually taken to show a prolonged failure in verb extension come from the set of studies by Imai and colleagues. Imai, Haryu, and Okada (2005), in a series of experiments with Japanese-learning preschoolers, demonstrated that 3-year-old children had trouble extending a novel verb to the same class of event with a different type of patient. In a forced-choice task, 3-year-olds, 5-year-olds and adults were familiarized with a scenario where an agent (e.g. a woman) performed some action (e.g. rolling) on a novel object (the patient) (e.g. a football-shaped orange Frisbee fins); those in the *noun condition* heard a descriptive sentence with a novel noun *X* (e.g. 'Mite! *X* ga aru', see (3-5) for linguistic gloss), and those in the *verb condition* heard a descriptive sentence with a novel verb *X*, (e.g. 'Mite! *X*-teiru!', see (3-6) for linguistic gloss). (Novel verbs were introduced in bare-verb sentences because Japanese is an argument-drop language.) At test, they were shown two novel scenes derived but different from the familiar one – in one scene, the agent performed the same action on a different type of novel object (e.g. a blue candle holder in the shape of a cup with a long stem); in the other, the agent performed a different action (e.g. tossing) on the same patient. Participants were asked to point to either of the scenarios with a prompt question – in the noun condition, the question was '*X* wa docchi ni aru?' (see (3-7) for linguistic gloss); in the verb condition, the question was '*X*-teiru no wa docchi kana?' (see (3-8) for linguistic gloss). Results were: in the noun condition, all three groups of participants pointed to the same-object-different-action scene significantly more than chance; in the verb condition, however, only adults and 5-year-old, but not 3-year-olds, pointed to the same-action-different-object scene more than chance. These results suggested that 3-year-olds seemed to have quite specific expectations about the newly learned verbs in this experiment – for example, *X* describes the set of rolling events performed on some particular type of patients, namely, the football-shaped-Frisbee-with-fins sort of things; and 5-year-olds' expectations seemed to be more general.

(3-5) Mite! *X*-ga-ar-u.

Look! *X*-NOM-be.inanimate-PRES

"Look! There's (an) *X*."

(3-6) Mite! *X*-teir-u.

Look! *X*-PROG-PRES

"Look! *X*-ing."

(3-7) *X*-wa-docchi-ni-ar-u?

X-topic-which-LOC-be.inanimate-PRES

"Which [movie] is *X* in?"

(3-8) *X*-teir-u-no-wa-docchi-ka-na?

X-PRES-PROG-nominalizer-TOP-which-Question.particle

"Which is the one who is *X*-ing?"[18]

The authors raised an alternative explanation for 3-year-olds' failure in verb extension in the above experiment: those children had the same general expectations as adults and their older peers, but they failed to demonstrate this knowledge in a task that might have overloaded their information-processing capacity. In the above experiment, to be able to make a choice between the two candidate scenes, children had to a) mentally process the event during familiarization, b) hold that event in working memory, c) process the two simultaneously presented candidate events at test, and d) retrieve the familiar event rom working memory to compare with the two candidate events. These steps may have gone beyond 3-year-olds' capacity. To reduce possible

[18] The linguistic glosses from (3-5) to (3-8) are based on the original paper's glosses as well as from consultation with Dr. S. Funakoshi (personal communication, April 7, 2015).

information-processing load, the authors conducted a second experiment with 3-year-old children, with a less-demanding yes-no paradigm. In this experiment, at test, instead of seeing two candidate events simultaneously, children saw them one at ta time, each presented with the familiar event next to it (to save them the effort to hold that event in working memory and later retrieve it). However, even with this methodological change, 3-year-olds were still not able to extend novel verbs to the same action with a different type of patient. Although this manipulation did not lead to 3-year-olds' success in verb extension, reducing the processing load the task imposes on child participants should definitely be taken into consideration.

In a more recent paper, Imai et al. (2008) extended the forced-choice task (the first experiment reviewed above) to another group of Japanese-learning children, as well as to English-learning preschoolers; in these experiments, they manipulated type of argument of the linguistic stimuli – sentences with overt arguments[19] (e.g. 'Look, she's *X*-ing it!'), and sentences with null arguments ('Look, *X*-ing!'). Results were: First, as a replication of Imai et al. (2005), Japanese-learning 5-year-olds, but not 3-year-olds succeeded in verb extension when the novel verbs were taught in null-argument sentences; but with full-argument sentences, even 5-year-old Japanese children failed (3-year-olds failed too). Second, 5-year-old English-learning children, but not 3-year-olds successfully extended a novel verb to same action with a different type of patient,

[19] The pronominal-vs-lexical-argument contrast was not considered in their study; they only focused on the overt-vs-null-argument contrast. So, in their study, the Japanese sentence with overt arguments used a full lexical NP (e.g. 'girl') whereas the English sentence with overt arguments used pronominal NPs (e.g. 'she', 'it').

but only in the full-argument condition. The difference between Japanese- and English-learning children's performances was attributed to the asymmetry between Japanese and English in their argument-drop status: null-argument sentences might be more natural to Japanese-learners and the extra information coming with overt-argument sentences might actually be a hindrance for them; to the contrary, English-learners might need overtly expressed arguments to support an accurate parse of the sentences. (See Imai et al. (2008) for detailed discussions on cross-linguistic differences[20].)

The upshot Imai and colleagues took from these results (Imai et al., 2005; Imai et al., 2008) was that failure in verb extension tasks was seen quite robustly in preschool-aged children learning different languages; and they took these results to suggest this: verb meanings are initially represented quite specifically, and young children do not seem to "appreciate the full meanings of the verbs"; "verb meanings develop progressively, and linguistic experience is necessary to finally obtain fully adult-like representations of verb meanings" (Imai et al., 2008); also see Gallivan (1988), Imai et al. (2005), and Theakston, Lieven, Pine, and Rowland (2002). The upshots I take from the same results, on the contrary, are: First, children successfully extended novel verbs to new instances with a different type of patient at least in some circumstances; this suggests that these children (i.e. 5-year-old English- and Japanese-learners) have the knowledge that verbs are relatively broadly extendible. Second, their performance was very much

[20] They also tested Chinese-learners, but the picture was more complex: since Chinese lacks inflectional morphology for a novel word presented in null-argument structures to be unambiguously identified as a verb, it's hard to interpret data from the null-argument condition.

contingent on details of the task - argument type of the linguistic stimuli in this case; this suggests that young learners' demonstration of some knowledge may be very sensitive to task designs, and that caution should be taken to separate lack of knowledge from failure of demonstrating the knowledge. In designing an appropriate task, grammatical, pragmatic, cognitive factors, the interactions among them, as well as language-specific properties, should all be taken into consideration, as much as possible. The above findings at least highlighted the role of argument type as a potential factor.

3.1.1.2 Some Cases of Success in Verb Extension

In contrast to the above reviewed failures, several cases of success in verb extension tasks were also seen in the literature. One such case were from a third experiment of Imai et al. (2005). In this experiment, the authors examined 3-year-olds' ability to extend novel verbs to the same action with a different *agent* (originally person A was the agent, and at test person B was the new agent), while the patient object was kept constant, and showed that 3-year-olds were able to do so. Similarly, Golinkoff, Jacquet, Hirsh-Pasek, and Nandakumar (1996) presented 3-year-olds (average 34 months) pictures depicting some actions performed by some Sesame Street character; in one condition, children were shown two unfamiliar actions, one just-learned with a novel verb as its label (e.g. *daxing*) but performed by a different agent (a different Sesame Street character), and the other one a new action; and children were able to choose the just-learned action over the new action significantly more than chance[21]. Kersten and Smith (2002) demonstrated 3.5- to 4-year-olds were able to reject a same-agent-different-motion event as an example of a novel verb and accept a different-agent-same-motion one when the agents were familiar objects – for example, a car performing a motion versus a truck performing the same motion. All these results suggested children about 3-4-years of age were able to extend a newly learned verb to the same event type with a different agent.

Some other studies also reported success in children younger than 3 years. Forbes and Poulin-Dubois (1997), using a preferential looking task, presented 20- and 26-month-olds two familiar motion events – kicking and picking-up, performed by two female actors (person A kicking, person B picking up); then children were tested with these two events shown on opposite sides of the screen, each with a different agent (person C kicking, person D picking up), and were asked ‘where is kicking’ or ‘where is picking up’. They found that 26-month-olds, as well as 20-onth-olds with higher expressive vocabulary, consistently paid more attention to the target event than the foil, despite the change in agent identity. In another preferential-looking task, Maguire et al. (2002) tested 18-month-old infants. Infants were taught a novel intransitive verb (e.g. ‘she’s *blicking*’) concurrent with a scene where an agent (person A) performed an action (action 1); and when they were later shown on opposite sides of the screen a different-

[21] This study also tested whether children preferred to choose an unfamiliar action (i.e. one that they didn’t have a name for) as the referent of a novel verb; so the novel verb was taught in an implicitly: in one trial with three familiar actions (e.g. reading, sleeping, eating) and one unfamiliar action (e.g. doing an arabesque), children were asked ‘where’s *daxing*?’, and they successfully pointed to the unfamiliar action (e.g. doing an arabesque). This novel verb and the novel action it labels then became the ‘just-learned’ verb and action in the next trial, where verb extension ability was tested.

agent-different-action (person B, action 2) scene and a different-agent-same-action (person C, action 1) scene, they preferred the latter[22]. These all seem to suggest that the ability to extend a verb (familiar and novel) to the same event type with a different agent is present in infancy. For the purpose of the current chapter – to identify the learner's initial expectation about verb meaning, though, familiar verbs are not so informative, because it is hard to distinguish whether a relatively general expectation about verb meanings is learned from experience (e.g. induced from many instances of kickings and picking-ups performed by different agents) or is part of the learner's initial expectation.

Notice that in the studies reviewed above, failures seem come from extension to different *patients*, whereas successes seem to come from extension to different *agents*. Even in the same sets of studies, presumably using identical experimental settings, Imai et al. (2005) showed 3-year-olds succeeded in extending a novel verb to a different *agent* but failed with a different *patient*. The authors did not discuss what might have led to this asymmetry. Here I discuss two conjectures of mine.

One possibility may have something to do with an observation about a closer relation between verb meaning and the patient (or verb and internal argument), in comparison to verb meaning and the agent (or verb and external argument) (Kratzer, 1996, 2003; Marantz, 1984). It is often observed that "a

[22] They only did so when the agent actors during training was presented in point-light display were presented (Johansson, 1973) – a method that made the specific features of the actor herself less visible. Absent this particular of presentation, infants did have a preference of one scene over the other. The authors attributed this difference to infants' attention being easily captured by the object/person involved in an event.

particular kind of internal argument triggers a particular interpretation of the verb" (Kratzer, 1996); see (3-9) for an example of different interpretations of *throw* triggered by different kinds of internal argument (Marantz, 1984). However, the same is not often seen with external argument. It is possible this observation about the grammar is also reflected in children's early expectations: children may expect verbs to describe events that place specific semantic restrictions on internal arguments more than on external arguments, more willing to extend novel verbs to new agents than to new patients.

(3-9) a. throw a baseball

b. throw support behind a candidate

c. throw a boxing match (i.e. take a dive)

d. throw a party

e. throw a fit

Alternatively, perhaps the observed asymmetry is nothing deep but simply an artifact of task design. Notice that in Imai and colleagues' experiments testing extension to a different patient, not only the identity of the patient was different, but the new patient also seemed to be a *different sort of thing* – for example, it is fair to say that a football-shaped orange Frisbee fins and a blue candle holder in the shape of a cup with a long stem are two different types of objects; and it may be the case that young learners expect one type of object can be involved in the event labeled by the novel verb, whereas other sorts of objects cannot. On the other hand, in their experiment testing extension to a different agent, as well as all other studies reviewed above on agent extension, the original agent and the new agent were reasonably classified as the *same sort of thing* – either two different people (person A and person B), or two different vehicles (a car and a truck); and children expect entities of the same sort can be involved in the same type of event. Therefore, the asymmetry between children's performance in the agent- and the patient-test may actually reflect the asymmetry between extension to a different sort of participant and to a different participant of the same sort; the former possibly falls outside of children's expectation and the latter still within.

On the continuum of most specificity (S) to most generality (G), with respect to expectations about verb meanings, expecting a verb to apply to anything lies at the G end, expecting it to only apply to a particular instance of participant (but not other instances of the same sort) lies at the S end; and expecting a verb to apply to some sort of participants that share certain common features (e.g. animacy) lies somewhere in between. What determines the boundary of falling in or out of the same *sort* is not entirely clear – for example, one can say a car and a truck belong to the same sort of thing because they are both vehicles; or one can arguably say cars and trucks are different sorts of things). For the current purpose, at least we can try to decide if the learner has a more general expectation than that at the S end – whether young learners accept a verb to apply to the same event type with a different participant that arguably belongs to a different sort.

3.1.1.3 A Case of Success with Patient Extension in Infancy

Contrary to (Imai et al., 2005)'s finding that even 3-year-olds failed to extend a novel verb to the same event with a different sort of patient, below I review a study that shows successful *patient* extension with 24-month-old infants, in which the patient objects were arguably of two *different sorts* (e.g. a balloon sort of thing versus a rake sort of thing).

Waxman et al. (2009) tested 24-month-old English-learning infants' ability to extend a novel verb to a different sort of patient, while holding the agent constant. During *familiarization*, infants were presented with four consecutive exemplars of an actor performing an action on an inanimate object (e.g. a male character waving a balloon); in each exemplar the object was a different token of the same category (e.g. different balloons); and they heard a sentence like 'look, the man is *larping* a balloon' if in the verb condition, and 'look, the man is waving a *larp*' if in the noun condition, and 'look at this' if in the no-word condition. This was followed by a *contrast* phase, where an event of the same actor performing a different action on a different object (e.g. the male character playing a toy saxophone) was played, paired with an audio like 'he's not *larping* that', or 'that's not a *larp*', or 'look at that', contingent on the condition; then the original event came back on with the original audio; the role of the contrast phase was to emphasize that the correct application of the novel words was not unlimited. Finally at *test*, infants were shown two scenes on opposite sides of the screen – one a familiar scene, (e.g. the man waiving a balloon), the other is an unfamiliar scene with a change in the patient object (e.g. the man waving a rake);

these were first paired with a *baseline* prompt (e.g. 'now look, they're different'), and then infants were direct to look at the screen by a *response* prompt (e.g. 'which one is *larping*', 'which one is a *larp*' or 'what do you see now', depending on the condition). Results were: infants in all conditions showed a novelty preference during baseline; in the response window, this preference was overcome in the noun condition – attention shifted to the familiar scene, suggesting they mapped the novel noun to the object, rather than the action; but those in the verb condition and no-word condition maintained the baseline novelty preference. Infants' performance in the verb condition was inconsistent with them taking *larp* to only mean balloon-larping, because if they were, they would have shifted their attention to the familiar scene during the response window. Instead, their performance suggested that they successfully extended the newly learned verb to the same event type with a different sort of patient, thus waving a balloon and waving a rake were both instances of *larping*, giving no motivation for them to shift attention away from the novel scene to familiar scene.

Waxman et al. (2009) and Imai and colleagues' studies (Imai et al., 2005; Imai et al., 2008) both tested verb extension to a different sort of patient, but one showed success with 24-month-old infants whereas the other showed failure with 36-month-olds preschoolers. This contrast may not reflect a difference in young learners' knowledge, but the way experimenters probe it. In particular, the test children faced in Waxman et al. (2009) might be less demanding than that in Imai et al. (2005)[23] and Imai et al. (2008): in the former, children's choice was between a familiar scene and a novel one, which was presumably less demanding than choosing between two novel scenes as in the latter. Additionally, the use of multiple versions of a given scene (e.g. a man waving different balloons) and a contrast phase may each has a facilitative effect in word learning (Au & Markman, 1987; Clark, 1988; Hall & Bélanger, 2005; Waxman & Booth, 2003; Waxman & Markow, 1995).

3.1.2 Minimize Task Demands to Reveal Knowledge

I reviewed in the above section several studies in the literature reporting mixed results in young children's ability to extend a verb to the same event type with a changed participant, and alluded to a point I take to be very important in interpreting results of experimental linguistics. Here I repeat it: while caution should be taken in interpreting *successes* in laboratory experiments – to separate success as a true reflection of the grammatical knowledge of interest from success due to merely some task-accomplish strategy/heuristic, similarly, cautious should also be taken in interpreting failures – to separate true lack of grammatical knowledge from failure to demonstrate some existing knowledge.

One reason child participants may fail to demonstrate their grammatical knowledge is that the task overloads their cognitive capacity - processing load for

[23] As reviewed, they used a force-choice task as well as a yes-no task for 3-year-olds; the yes-no task was used to reduce information-processing load, but did not improve children's performance. Here, the comparison made here was based on their force-choice task.

instance. There are many ways young children may be overloaded. Extralinguistic information required to be processed, kept in memory and placed into comparison may be one, as pointed out by Imai et al. (2005). Linguistic information may be another, as evidenced by Lidz et al. (2009).

Lidz et al. (2009) showed that 22-month-old infants learned the meaning of a novel verb better from a sentence with a pronominal subject (e.g. 'it's *blicking*') than from a sentence with a full lexical NP as subject (e.g. 'the flower's *blicking*'). They thus hypothesized: although a lexical NP carries more information than a pronoun, in the sense of unambiguously identifying the referent, this informativity may come at a cost of extra processing load. They further tested this hypothesis in a second experiment: they added a noun familiarization phase at the beginning where infants saw the object (e.g. the flower) and heard audio stimuli naming the object several times (e.g. 'Hey, it's a flower; there's the flower again; did you see the flower'); this was supposed to reduce the processing burden associated with the lexical NP subject; and results were, as predicted, infants in the lexical NP condition succeeded in learning the novel verb this time. These findings suggested that a) these infants were able to learn the meaning of the novel verb in the task, b) their failure was not an indicator of lack of grammatical knowledge, but simple failure of knowledge demonstration, and c) reducing processing load imposed by the subject NP helped reveal their knowledge.

Therefore, young infants seem to be particularly vulnerable to information processing load such that they fail to demonstrate their true linguistic knowledge once they are overloaded. Lidz et al. (2009) demonstrated their vulnerability to the extra information brought by lexical NPs and their subsequent failure in verb learning task, in contrast to their success with pronominal NPs. We also manipulate subject type in the current study, to reduce processing load (with pronominal subject) as well as to further test Lidz et al. (2009)'s hypothesis in a verb extension task.

3.2 The Current Study

3.2.1 Hypotheses

Despite failures of verb extension in some cases (Imai et al., 2005; Imai et al., 2008 - 3yo, and 5yo given certain argument types), there were circumstances where young learners demonstrated willingness to extend verb meanings to different agents (Forbes & Poulin-Dubois, 1997 - 20mo and 26mo; Golinkoff et al., 1996 - 3yo; Imai et al., 2005 - 3yo; Kersten & Smith, 2002 - 3.5yo-4yo; Maguire et al., 2002 - 18mo), and to different patients (Imai et al., 2005 - 5yo given certain argument types; Waxman et al., 2009 - 24mo). Though the cases of success varied in subject ages, event participants (agent or patient), scope of specificity (a different thing or a different sort of thing), etc., they establish some grounds for the following hypotheses of mine: A. young infants, hold an initially relatively general expectation about verb meanings, and this general expectation allows them to extend newly learned verbs to the same event types with a different sort of participant, and B. finer semantic constraints different verbs place on their arguments are later learned and gradually added to the developing lexicon.

These hypotheses are different in two important ways from Imai and colleagues' opinion that verb meanings are initially quite specific and develop progressively into fully adult-like representations (Gallivan, 1988; Imai et al., 2005; Imai et al., 2008; Theakston et al., 2002). First, we disagree in the degree of specificity/generality of the learner's initial expectation about verb meanings – more generality is endorsed by the current hypothesis. Second, we disagree in what is meant by "fully adult-like" – given that verbs vary in how much specific semantic restrictions placed on their arguments (e.g. *gallop* vs. *fall*) even in adult representations of verb meanings, it is hard to say whether "fully adult-like" more generality or specificity; the current hypothesis implies a certain continuity between child learners' and adults' expectations, for both of whom a novel verb is expected to be broadly extendible, until more evidence about its semantic restrictions is given (e.g. a novel verb is seen to only apply to animals). The current study is not committed to testing the developmental aspect of this hypothesis (i.e. B)), but focuses on testing the initial stage of development (i.e. A).

To test this hypothesis, the current study examined *23-month-olds'* ability to extend newly learned intransitive verbs to a different sort of *agent*; and verbs were taught in sentences with different *types of subject NPs*. We tested 23-month-olds, because infants of 23 months are considered verb learning beginners, provided that infants' productive vocabulary does not start to include lots of verbs until the end of the second year (Bates et al., 1994; Caselli et al., 1995; Fenson et al., 1994; Gentner, 1982; Gleitman et al., 2005). We examined agent extension for two reasons: a) all previous studies on agent extension, to our knowledge, tested extension to a different instance of an agent or the same sort (e.g. person A and person B), which is a rather specific expectation, and we would like to test whether the learner's expectation is more general that enables them to extend to a different *sort of* agent; and b) if the observation that the internal argument bears a closer relationship than the external argument to the meaning of the verb (Kratzer, 1996, 2003; Marantz, 1984) is also reflected in the learner's initial expectation (see Section 3.1.1.2), it is possible that there is some asymmetry between extension to different agents and to different patients, and for this reason, testing extension to agents seem to be a better place to start with. We taught infants new verbs in sentences with different types of subject NPs, to fulfill a secondary goal – to test the hypothesis that the extra information (about referent identify) brought about by lexical NP subjects may actually overload infants' processing capacity and prevent them from showing their true linguistic knowledge which would otherwise be revealed with pronominal NPs as subjects (Lidz et al., 2009).

To sum up, the current study tests two hypotheses: first, the hypothesis that young infants at the beginning stage of verb learning are able to extend a newly learned verb to the same event category with a different sort of agent; and second, the hypothesis that having a pronominal subject better facilitates infants' verb extension than having a lexical NP subject, by virtue of imposing less processing burden.

3.2.2 Design Overlook

Following Waxman et al. (2009), Maguire et al. (2002) and Forbes and Poulin-Dubois (1997), we adopted a variant of the Preferential Looking Paradigm. This paradigm was first introduced into the literature by Golinkoff, Hirsh-Pasek, Cauley, and Gordon (1987); it is based on the well-tested assumption that young children prefer to look at a test image/scene that matches the auditory stimulus, than one that does not match. This method is known for its relatively low demands on the learner's part - requiring only visual fixation, and has been proven to be an effective way to probe young children's language comprehension: from infants as young as 2 to 6 months (Baier, Idsardi, & Lidz, 2007; Kuhl & Meltzoff, 1982; Tincoff & Jusczyk, 1999) to as old as 3 years (Naigles, 1998). In our adaptation of this paradigm, we taught the infant participant a novel intransitive verb, as a label for a concurrent event of an object performing a self-propelled action (e.g. a truck moving back and forth), and then presented the infant two novel scenes on opposite sides of the screen – in both, the original object was replaced by a novel object (e.g. a cart); in one, the novel object performed the familiar action (e.g. the cart moving back and forth), whereas in the other, the novel object performed a new action (e.g. the cart falling). Infants' preference to the familiar action during test more than during a predetermined baseline was taken as evidence that they successfully extended the meaning of the newly learned verb to an instance of the same event type but with a different sort of agent.

In this experiment, we took extra caution in making the task less demanding, to give infants a better chance to demonstrate their true linguistic knowledge. Here are the specific steps taken.

First, we used familiar objects (e.g. truck, ring tower, door, flower) during familiarization, and also included an object demonstration phase, to minimize the chance that their attention was not overly captured by the objects when they were taught the novel verbs (for previous literature that reported such object/person attraction effect, see Behrend, 1990; Brown, 1957; Kersten & Smith, 2002; Maguire et al., 2002).

Second, we included a contrast phase after familiarization, following Waxman et al. (2009), to demonstrate that there are limits on the application of the novel label, which has been shown to be facilitative in word learning (Au & Markman, 1987; Clark, 1988; Waxman & Booth, 2003; Waxman & Markow, 1995).

Third, in choosing the two candidate scenes, to test verb extension, it is necessary to have one scene presenting a new object and the familiar action: suppose we have action A with object A in training, it is necessary to have action A with object B as one test scene (i.e. the target). This is what every such experiment had; the difference between experiments lies in the choices of the other candidate scene. In Waxman et al. (2009), to reduce processing load, they used a same-object-same-action scene (i.e. action A with object A) as the other candidate, one that added no extra information to process. This design, however, may have a drawback in obtaining compelling data: both scenes at test are consistent with the novel verb's meaning (because both had the same action, action A); therefore, the best we could expect from infants, even if they do expect verbs to be broadly extendible, is that they begin with a novelty preference during baseline, and stay there during test, because they do not have strong motivation to shift their attention away from the interesting novel scene, provided the novel scene is also consistent with the auditory direction (e.g. 'which one is *daxing*?'). And in fact, this was what they observed; and such data, although consistent with the successful-extension hypothesis, were not by itself compelling, not without comparison to the other conditions. Therefore, in our design, we would like the other candidate scene to be *a true distractor*, rather than another target; this meant the other scene would need to present *a new action* (i.e. action B). We still had two options: a) action B with object A (same object as in training); and b) action B with object B (different object from that in training, but same object as that in the target scene). Option a) is what was adopted in Imai and colleagues' studies (Imai et al., 2005; Imai et al., 2008), but we took option b), trying to minimize possible demands imposed on making comparisons between the two candidate scenes: given the target scene is action A with object B, the distractor in option b) is minimally different from the target (only the actions are different), whereas the distractor in option a) differs form the target in both the action and the object, possibly imposing a higher demand.

Last but not least, we manipulated the type of syntactic subject used in the linguistic stimuli – a) a pronominal subject (e.g. '<u>it</u>'s *krading*'), b) a lexical NP subject (e.g. 'the truck's *krading*'), and c) an NP subject that does not carry much lexical content (e.g. 'that thing's *krading*'). Having the pronominal-subject condition (i.e. *it*-condition), based on Lidz et al. (2009) was to reduce possible processing load coming from the linguistic stimuli; having the lexical-NP-subject condition (i.e. *NP*-condition) was to test our secondary hypothesis - whether having a pronominal subject, would better reveal learner's knowledge than having a lexical NP subject. And the non-lexical-NP-subject condition (i.e. *thing*-condition) was included as a control: if infants' performance in the *it*- and *NP*-conditions did differ, whether that difference came from a difference in lexical content or lexical category. In particular, if infants' performance in the *thing*-condition patterned more closely with that in the *it*-condition, then it was probably lexical content that mattered, because both *it* and *thing* are low in lexical information; if performance in the *thing*-condition patterned more closely with that in the *NP*-condition, then it was likely that lexical category that mattered.

3.3 The Experiment

3.3.1 Participants

Eight-three English-speaking infants (37 boys, 39 girls) with a mean age of 22;28 (range: 20;03 – 26;12) participated in this experiment. Two additional infants were tested but excluded from the final sample because of being unable to finish the experiment. All infants were recruited through the Infant Studies Consortium Database at University of Maryland College Park.

3.3.2 Stimuli

The visual stimuli were videotaped events where some real-world objects engaged in some self-propelled actions; for example, a truck moved back and forth. The auditory stimuli consisted of recordings of a female native speaker of English producing sentences in child-directed intonation. The recordings were edited by hand for duration and timing and were then synchronized with the visual stimuli. The visual and auditory stimuli were digitized to a single source to ensure consistent presentation of stimuli across participants and experimenters. Each participant was presented two training trials, each lasting about 20 seconds, which followed the sequence described in Table 3.1(a); and four test trials, each lasting about 72 seconds, which all followed the sequence described in Table 3.2(a). See Table 3.1(b) for all stimuli used in training trials, and Table 3.2(b) for all stimuli used in test trials. The entire video was about 5.5 minutes.

Phase	**Visual Scene**		**Position on the screen**	**Audio track**
centering	baby smiling face		center	giggling
test - baseline	ball bouncing	ball swinging	on either side of the screen	oh, look, they're different.
test - response	ball bouncing	ball swinging		which one is bouncing?

(a)

Candidate events (left and right counterbalanced)		Familiar verb
ball bouncing	ball swinging	bouncing
cabinet twisting	cabinet opening	opening

(b)

Table 3.1: Summary of training stimuli – (a) sequence of training trials; () all stimuli used in training trials

Phase	Visual Scene		Postion on the screen	Audio track
centering	baby smiling face		center	baby giggling
object demonstration	truck (motionless)		4 presentations, each on a different quardrant of the screen, order counterbalanced across trials	Do you see it / the truck / that thing?
familiarization	truck moving back and forth		3 presentations, each on either side of the screen, order counterbalanced across trials	Look, it / the truck / that thing is *krading*.
contrast	truck rotating		center	Oh, no, now it / the truck / that thing is not *krading*.
	truck moving back and forth		center	Yay, now it / the truck / that thing is *krading*.
test - baseline	cart moving back and forth	cart falling on its side	one on the left, one on the right, order counterbalanced across trials	oh, look, they're different.
test - response	cart moving back and forth	cart falling on its side		which one is *krading*?

(a)

novel verb	**familiarization event**	**contrast event**	**candidate events at test (left and right counterbalanced across trials)**	
krade	truck moving back and forth	truck rotating	cart moving back and forth	cart falling
skain	ring lifting	ring circling	lava lamp lifting	lava lamp circling
doff	flower rotating	flower spinning	sprinkler rotating	sprinkler going up and down
flurb	door closing	door lifting	house closing	house rocking

(b)

Table 3.2: Summary of test stimuli – (a) sequence of test trials; (b) all stimuli used in test trials

3.3.3 Apparatus

The stimuli were played on a Samsung wall-mounted 51-inch plasma television, with built-in speakers, located 66 inches away from the chair (or highchair) where the infants were seated. A Sony EVI-D100 video camera was placed directly above the TV monitor. The experimenter observes the infant from another room, and will adjust the camera (by zooming in or out) to make sure the child's face is always centered, for the convenience of coding that will be conducted after the experiment offline. The video of the child, with a picture-in-picture display of what was on the TV screen, was captured on an iMac computer using QuickTime.

3.3.4 Design

This experiment has three conditions, varying in the type of subject the linguistic stimuli used: a) the pronominal-subject condition (henceforth *it*-condition), where every mention of the object in the scene was with the pronoun *it*; crucially, the introduction of the novel verb was in a sentence with *it* as the subject (e.g. 'it's *krading*'); b) the lexical-NP-subject condition (henceforth *NP*-condition), where every mention of the object was with its name (e.g. the truck); crucially, the introduction of the novel verb was in a sentence with an NP as the subject (e.g. 'the truck is *krading*'); and c) the non-lexical-NP-subject condition (henceforth *thing*-condition), where every mention of the object was with the noun phrase *that thing*; crucially, the introduction of the novel verb was in a sentence with *that thing* as the subject (e.g. 'that thing was *krading*'). For each condition, there were two orders, differing in the relative order of the four test trials; the two training trials did not vary in order across conditions/orders. See Table 3.3 for a summary.

	training1	**training2**	**test1**	**test2**	**test3**	**test4**
***it*, order 1**	bounce	open	*krade*	*skain*	*doff*	*flurb*
***it*, order 2**	bounce	open	*flurb*	*doff*	*skain*	*krade*
***NP*, order 1**	bounce	open	*krade*	*skain*	*doff*	*flurb*
***NP*, order 2**	bounce	open	*flurb*	*doff*	*skain*	*krade*
***thing*, order 1**	bounce	open	*krade*	*skain*	*doff*	*flurb*
***thing*, order 2**	bounce	open	*flurb*	*doff*	*skain*	*krade*

Table 3.3: Trial sequences across conditions/orders

3.3.5 Procedure

The procedure began with obtaining the parent(s)' informed consent and collecting the MacArhur Communicative Development Inventory (MCDI) (Dale & Fenson, 1996) – a standardized measurement of productive vocabulary development. When the infant was ready, he/she was led to the test room where the TV monitor and the digital camera were located. The parent came to the test room with the infant and stayed with him/her during the entire process. The infant sat either in the parent's lap or in a highchair in front of the monitor. We took precautions to ensure that the parent could not influence the child's behavior, by explicitly instructing the parent not to direct the infant's attention in any way, and asking the parent to wear a visor (to block sight) in cases where she chose to hold the infants on her lap.

The experimenter began the experiment in the control room next door, by setting up the computer to display the pre-recorded stimuli and to record the participant's attention. Each participant first received two *training trials*. Each training trial began with a display of a smiling baby face paired with baby giggling sound, presented at the center of the screen; the purpose was to capture and center the participant's attention. Then, the infant saw two events on opposite sides of the screen – in one, a familiar object performed an action (e.g. ball bouncing); in the other, the same object performed a different action (e.g. ball swinging). Concurrent with this display, the infant first heard a *baseline prompt* that was a non-directing audio (e.g. 'look, they're different'), and then he/she heard a *response prompt* that was a directing audio (e.g. 'which one is bouncing') to which only one of the events matched. These two training trials were to get the infant familiar with the task. The target of the first training trial was on the left, and the target of the second training trial was on the right. All infants in all conditions received the same training trials in the same order.

After that, each participant received four *test trials*. Each test trial also began with a display of the smiling baby face, to center the participant's attention. After that was three phases. First was the ***object demonstration phase***, where a real-world object in motionless state was shown, paired with an audio introducing it; the audio varied across conditions – e.g. 'Do you see it?', or 'Do you see the truck?', or 'Do you see that thing?'; this object introduction was presented four times, each time at a different quadrant of the screen, counterbalanced across trials. The purpose of this phase was to familiarize the infant with the object, to reduce the possibility of their attention being over-captured by the object when they were later taught the novel verb. Following this phase was the ***familiarization phase***, in which the demonstrated object was engaged in some self-propelled action (e.g. truck moving back and forth), paired with a linguistic stimulus containing a novel verb; the linguistic stimulus was an intransitive sentence with different types of subject, depending on the condition - e.g. 'it's *krading*', 'the truck is *krading*', and 'that thing is *krading*'. This was presented three times, each on one side of the screen (Left-Right-Left, or Right-Left-Right, counterbalanced across trials). Next came the ***contrast phase***, during which the same object engaging in a different action was presented (e.g. truck spinning), paired with a negation sentence (e.g. 'it's not *krading*', 'the truck is not *krading*', ‘that thing is not *krading*’); then the original event came back on (e.g. truck moving back and forth) paired with the a non-negation sentence (e.g. ‘it’s *krading*’, ‘the truck is *krading*’, and ‘that thing is *krading*’); events was presented on the center of the screen during this phase. Subsequently, the ***test phase*** began, where two new events were shown on opposite sides of the screen – in both, the agent was a new object (e.g. a cart, rather than the truck); in one event, the new object performed the familiar action (e.g. cart moving back and forth), whereas in the other, the new object performed a new action (e.g. cart falling on its side). The test phase was divided into two periods, in both of which the visual presentation was the same – two candidate events on two sides of the screen: first was the ***baseline period***, during which a non-directing audio was paired with the visual display (e.g. ‘look, they’re different’); this was to establish the infant’s baseline preference; following was the ***response period***, during which a directing audio was presented (e.g. ‘which one is *krading*’) to which only one of the candidate events was the match; the matching event was therefore the *target*, and the other event was the *distractor*. The side of the screen where the target was located was counterbalanced across trials.

3.3.6 Coding

The videotaped films of infants’ attention were digitized into a format where their eye movements could be coded on a frame-by-frame basis. A trained coder coded the film in Supercoder – a custom program for coding preferential looking videos.

The sound track was removed when coding was conducted, to ensure that the coder were blind to the target-distractor positions and to condition assignment. The coder identified for each frame (30 frames per second) whether the infant's eyes were oriented to the left scene, right scene, or neither, by pressing a different key corresponding to each type of look.

3.3.7 Measurement

The dependent variable commonly used in the Preferential Looking Paradigm (PLP) is the *average proportion of looks* towards the target within *a selected window*. The selection of such a window may vary from study to study. The traditional analysis Golinkoff, Hirsh-Pasek, Mervis, Frawley, and Parillo (1995) used was a 6-second window starting from the offset of the target word (i.e. disambiguation point). This analysis, however, was questioned by recent research of Fernald and colleagues on real-time measures of the time-course of children's gaze patterns in response to certain linguistic stimuli (Fernald, Zangl, Portillo, & Marchman, 2008), which found a sign of rapid processing towards the end of the second year of age: In several studies using a variation of PLP examining infants responses to familiar words as a function of different prosodic features, they found that 24-month-olds performed worse than 18-month-old when their look time was average over a 6-second window, countering the predicted improvement in word recognition; when they used a window of 4 seconds, however, the data begun to look more towards the predicted pattern; and eventually, a window of 2 seconds revealed a large proportion of look towards the target (80%) for 24-month-old and a smaller proportion (60%) for 18-month-olds. This manipulation in window selection showed that the traditionally used window may underestimate older infants' real knowledge by failing to capture their correct and yet fast-fading responses due to rapid processing. Therefore, a window of 2 seconds has now been adopted for most of PLP studies for children around or above 2 years of age; and using a window from the *onset* (rather than offset) of the disambiguation point is another manipulation aiming to capture children's rapid responses.

In this experiment, we adopted the well-tested 2-second window as our way of data reduction. Following Waxman et al. (2009) and Lidz et al. (2009), we selected two 2-second windows: a baseline window - one from the baseline period of the test phase, and a response window – one from the onset of the novel verb of the response prompt (e.g. 'which one is *krading*') that ended 2 seconds later. Within each window, we calculated for each infant and each trial the average proportion of looking time devoted to the target scene – total looking time devoted to the target, divided by the total looking time devoted to the target and the distractor. The value of this variable would always fall in the range of [0, 1]: a value equal to 0.5 would indicate the same amount of time devoted to the target and the distractor; a value above 0.5 would indicate more time spent looking to the target, and a value below 0.5 more time spent looking to the distractor[24]. We then calculated the mean for each infant across all test trials. Thus, for each infant

[24] This is because in data reduction, we coded a look towards the distractor as '0', and a look towards the target as '1',

participant, we obtained two values: average proportion of looking time to the target during baseline window, and average proportion of looking time to the target during response window. For convenience, I will refer to them as *baseline-target-preference* and *response-target-preference*.

3.3.8 Predictions

For the primary research question – whether infants are able to extend novel verbs to the same event category with a different sort of agent, the predictions were: if they were to able to do so, they would demonstrate a significantly larger response-target-preference than their baseline-target-preference. In addition, following Waxman et al. (2009) and Lidz et al. (2009), we may also expect to see a novelty preference during baseline. Although both scenes at test were novel, because both contained a novel agent, but the one that additionally involved a novel action probably carried more novelty. If so, then we would expect infants started with preference to the distractor during baseline, and shifted their attention away from the distractor and more towards the target during responses, *if* they learned the meaning of the novel verb and expected verbs to be relatively broadly extendible.

For the secondary question – whether the type of subject in the linguistic stimuli affects infants' verb learning in a verb extension task differently; in particular, whether novel verbs taught in sentences with pronominal subjects were better learned and extended than those taught in sentences with lexical NP subjects. Following (Lidz et al., 2009), we may see successful verb extension in the *it*-condition, but not *NP*-condition; or, a larger magnitude of learning in the *it*-condition than the *NP*-condition if infants succeeded in both. As for the *thing*-condition, it would help to distinguish between two explanations in face of a predicted difference between *it*- and *NP*-condition: if performance in the *thing*-condition patterned together with that in the *it*-condition, then what drove the distinction between *it*- and *NP*-condition was likely to be a difference in their lexical context, thus a different in information-processing load; if performance in the *thing*-condition patterned together with that in the *NP*-condition, then it was more likely that it was lexical category that mattered, rather than lexical content.

3.3.9 Results

Infants' looking time data were entered into a three-way mixed ANOVA with *condition* (it vs. NP vs. thing) and *order* (order 1 vs. order 2) as between-participant factors, and with *window* (baseline vs. response) as a within-participant factor. This analysis revealed a main effect of window, $F(1, 70) = 5.20$, $p = 0.03$; and a significant interaction between condition and window, $F(2, 70) = 3.65$, $p = 0.03$; no other main effect or interaction was revealed. Since there was no main effect of order – $F(1, 70) = 3.23$, $p = 0.08$, data from order 1 and order 2 were subsequently collapsed.

The collapsed data were then entered into a two-way mixed ANOVA with *condition* as a between-participant factor, and *window* as a within-participant factor. This analysis revealed a main effect of window, $F(1, 73) = 5.35$, $p = 0.02$;

and a significant interaction between condition and window, $F(2, 73) = 3.75$, $p = 0.03$. Based on these, for each condition, we conducted planned comparison between baseline and response windows, testing the alternative hypothesis that response has a greater target preference than baseline. Specifically, for the *it*-condition, response-target-preference ($M = 0.58$, $SD = 0.25$) was significantly greater than baseline-target-preference ($M = 0.42$, $SD = 0.13$), $t(24) = -2.79$, $p < 0.01$; for the *NP*-condition, response-target-preference ($M = 0.46$, $SD = 0.23$) was greater than baseline-target-preference ($M = 0.50$, $SD = 0.18$), but not significantly greater, $t(24) = 0.78$, $p = 0.78$; for the *thing*-condition, response-target-preference ($M = 0.51$, $SD = 0.19$) was significantly greater than baseline-target-preference ($M = 0.42$, $SD = 0.13$), $t(24) = -1.89$, $p = 0.04$. See Figure 3.1 for illustrations of the results. It is also important to notice that in both *it*- and *thing*-conditions, looking proportions to the target during the baseline window were significantly below chance: *it*-condition, $M = 0.42$, $p < 0.01$; *thing*-condition, $M = 0.42$, $p < 0.01$. However, in noun-condition, looking proportions to the target at baseline were not below chance: $M = 0.50$, $p = 0.47$.

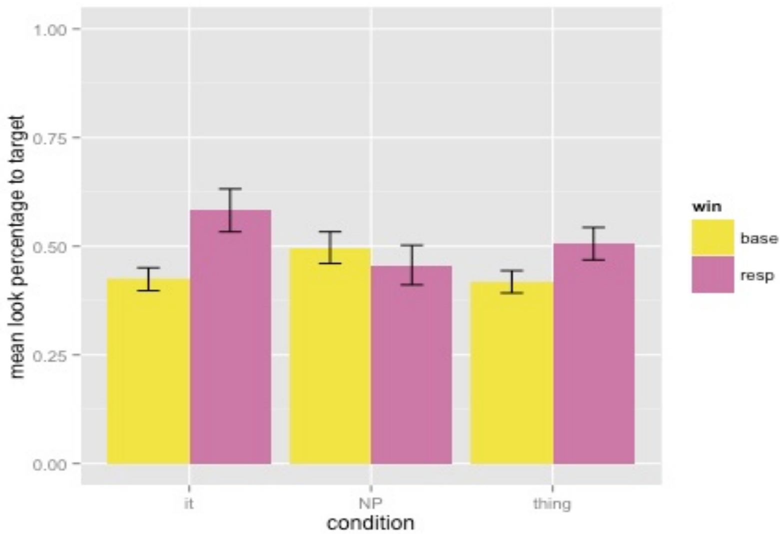


Figure 3.1: Mean look proportion in each window across conditions

3.3.10 Discussion

These results suggested 22-month-old infants were able to extend novel verb meanings to the same event type with a different sort of agent, at least under some circumstances. In our experiment, their success was conditioned on the type of subject in the linguistic stimulus they received. In particular, when they were taught the novel verbs in sentences with pronominal subjects (e.g. 'it's *krading*') or with the NP 'that thing' as subjects (e.g. 'that thing is *krading*'), they successfully overcame their novelty preference at baseline and shifted their attention towards the less novel yet correct scene at response. Their success in these conditions suggested that infants this age hold the expectation that verb meanings are relatively general and thus broadly extendible.

Nevertheless, these infants failed in the *NP*-condition where they learned the novel verbs in sentences with a lexical NP as subjects (e.g. 'the truck is *krading*'). The asymmetry between the *it*-condition and the *NP*-condition replicated Lidz et al. (2009)'s findings. And the observation that infants' performance in the *thing*-condition patterned with that in the *it*-condition, rather than with that in the *NP*-condition, suggested that what was driving the *it-NP* asymmetry was more likely to be a difference in their lexical content, rather than a difference in their lexical category. This is because in terms of lexical category, '*that thing*' and NPs like *'the truck'* fall into one category – determiner/demonstrative + noun, whereas '*it*' is pronominal; but in terms of lexical content, both '*that thing*' and '*it*' carry little information about the referent picked out, whereas NPs like '*the truck*' specifies a referent object in the world. As Lidz et al. (2009) suggested, although more lexical content provides more information, recognizing and utilizing that information may sometimes come at a cost.

3.4 General Discussion

To briefly summarize, the current study presented one experiment using an adaptation of the Preferential Looking Paradigm, with special care paid to reduce possible barriers preventing infants from demonstrating their linguistic knowledge, which demonstrated that English-learning 22-month-old infants were able to learn a novel verb and extend it to the same event category with a different sort of agent.

A manipulation with the types of subject in the linguistic stimuli, which potentially varied in the amount of information carried, further revealed that these young infants were particularly vulnerable to cognitive demands, and that they only successfully demonstrated their verb extension ability in conditions where information-processing demands coming from the linguistic stimuli were lower. In what follows, I will discuss the implications of these findings (Section 3.4.1), some interesting observations and potential caveats from the results (Section 3.4.2), some directions for future work with regard to the central question at issue (Section 3.4.3), as well as some thoughts on related issues (Section 3.4.4).

3.4.1 Implications

I discussed in Chapter 2 that the learner expects verbs to described kinds of events (i.e. the verb-event bias). This chapter raises a further question: what kinds of event concepts does the learner expect verbs to describe? Verbs in general all describe the ways in which participants are involved in events; on top of this, does the learner also expect verbs to *additionally* describe what kinds of things the participants are? For instance, does the learner expect a novel verb *gorp* heard in an instance of a cat involved in event *e* to only describe events with cat-sort-of-thing as a participant (i.e. more specific), or does she expect *gorp* to be applicable to other types of participants s as well (i.e. more general)? Verbs vary in their specificity/generality with regard to how much semantic restrictions placed on their arguments (e.g. *gallop*, requiring its participant to be horse-sort-of things, is more specific than *spin*). Of particular interest here is: despite these variations that are presumably learned on a verb-by-verb basis, what is the learner's *initial* expectation about the specificity/generality of a newly learned verb. Regarding this question, we hypothesize that the learner's initial expectation of verb meanings is relatively general, with finer verb-specific semantic restrictions later added to the developing lexicon.

This question is usually evaluated in laboratory by means of testing young learners' ability to extend a newly learned verb to the same event type with a different (sort of) participant; for example, if a novel verb *gleeb* is learned in an instance of action A with participant A involved, we could ask if young learners are willing to accept another instance of action A with participant B involved as a correct exemplar of the verb *gleeb*. In such tasks, successful extension is usually taken as evidence for a more general expectation about verb meanings, whereas failure to extend as evidence for a more specific expectation. A series of studies from Imai and colleagues (Imai et al., 2005; Imai et al., 2008) on several different languages reported several failures in preschool-aged children, and they took these results to suggest an initially rather specific representation of verb meanings, which would require linguistic experience to gradually become more general. We suggest interpreting the reported failures with caution, because they may not reflect lack of knowledge, but simply reflect a failure in demonstrating the knowledge due to task demands. In fact, in these studies, successful extension was observed under some circumstance; for example, when novel verbs were taught in bared-verb sentences (i.e. null arguments) for Japanese 5-year-olds, and in full-argument sentences for English 5-year-olds, these children succeeded; for another example, when the extension task was about a different agent rather than a different sort of patient, 3-year-olds succeeded. In addition, several other studies reported success in verb extension tasks under some circumstances as well (Forbes & Poulin-Dubois, 1997; Golinkoff et al., 1996; Kersten & Smith, 2002; Maguire et al., 2002; Waxman et al., 2009).

Most of these reported successes, however, do not actually answer our question, because some of them tested children beyond the initial stage of verb learning (e.g. 3~5-year-olds), and some of them tested verb extension to a different instance of a participant of the same sort (e.g. person A versus person B, both the human sort) rather than extension to a different *sort of* participant (e.g. cart-sort-of-thing versus flower-sort-of-thing). For our purpose, we are interested in young infants who are at the initial stage of verb learning, and their ability to extend novel verbs to a different *sort of* participant – more general than extension to merely a different instance of participant of the same sort. Waxman et al. (2009) was the only one, to our knowledge, that showed successful verb extension to a different sort of participant (i.e. patient) in young infants (i.e. 24-month-olds). None has showed extension to a different sort of *agent* in younger infants, and the current study closes this gap.

The current study paid efforts to reduce possible cognitive demands that could prevent infants from demonstrating their real linguistic knowledge. One particular manipulation was the type of subject NPs, in light of Lidz et al. (2009). Two experiments in Lidz et al. (2009) showed that 23-month-olds learned novel verbs better from sentences with a pronominal subject than from sentences with a lexical NP subject, and that the advantage of pronominal subject should be attributed to less lexical content to process, rather than other reasons like providing a better context for verb categorization (e.g. *it's _ing* is a frequent frame for verb). Following them, the current study hypothesized a cue with more information may sometimes come at a cost of processing load, interference with infants' knowledge demonstration.

The findings of the current study support our hypotheses. With respect to the hypothesis about the primary question - that infants' initial expectation about verb meanings is relatively general, we showed 22-month-olds were able to extend newly learned verbs to a different sort of agent, lending support for the generality of their initial expectation. With respect to the secondary hypothesis about the role of subject NPs, we showed these infants successfully demonstrated their knowledge when information processing load was lower (i.e. in *it* and *thing* conditions) but failed to do so with more information to process (i.e. in *NP* condition).

3.4.2 Some Observations and Caveats

One observation in the current experiment does not align with Lidz et al. (2009)'s findings. In Lidz et al. (2009), consistent novelty preference during baseline was observed; but in the current experiment, no novelty preference was observed in the NP condition. This is probably because in Lidz et al. (2009), the two candidate events at test varied distinctively in novelty – one was familiar, one was novel; in the current experiment, however, both events were novel, in terms of both including a new object, but they only differed in relative novelty – the one with a new action was presumably more novel than the one with a familiar action. This said, it remains a puzzle why a novelty preference was observed in *it*- and *thing*-conditions, but not in the *NP*-condition. One possible explanation is: the NP-condition, by virtue of mentioning the referent object explicitly, may have drawn attention to the familiar object, thus making a change in the identify of object particularly capturing; as a consequence, infants equally preferred the two test scenes, for both of them involved a novel object. No matter what the real explanation is, the differences across conditions were prominent, both in terms of baseline preference, and in terms of attention shift from baseline to response.

Another issue is probably more concerning. In the second experiment of Lidz et al. (2009), when processing load was reduced by having a noun-familiarization phase at the beginning, infants in the NP-condition also succeeded. In our experiment, during the object demonstration phase, in the NP-condition, the object was labeled several times by the NP (e.g. 'do you see the truck?'). If this manipulation successfully reduced processing load for infants in Lidz et al. (2009)'s experiment, why did it fail to do so for infants in the current experiment? This is something we don't have a better answer for, other than that a verb extension task is presumably harder than a verb learning task, and that infants are particularly vulnerable to processing loads in a harder task.

One last issue I would like to raise is an alternative explanation for infants' success in the *it*- and *thing*-conditions. Infants could have used a strategy that said ‘find a scene that had some common part as the scene in familiarization’, and the only scene that matches this strategy was the target (novel object, familiar action). However, in using such a strategy, they would have to be able to isolate the event (i.e. the action in this case) from its participant in their event representation; if they were able to do so, it seemed unlikely that they were not able to entertain a verb meaning extendible to different participants. Unlikely as it appears, this alternative still holds and deserves further examination.

3.4.3 Future Directions

The findings of the current study showed that 22-month-old English-learning infants were able to demonstrate their ability of extending a newly learned verb to the same event category with a different sort of agent, under conditions that imposed less demands on their other cognitive abilities. I think these findings are revealing about the expectations the language learner holds at the outset of verb learning – the learner may expect verbs to describe *the way* in which the participants are *involved in* events, but does not expect verbs to also describe what *kinds of things* the participants are; in other words, the learner entertains a relatively general expectation about verb meanings, allowing newly learned verbs to be relatively broadly extendible. In addition to current findings, I believe more compelling support for this argument should additionally receive evidence from at least the following two aspects.

First, this argument would be more convincing if more evidence is seen from the *patient* extension case, where most controversy lies. The only evidence of successful extension, to our knowledge, comes from Waxman et al. (2009)'s findings with 24-month-olds. Imai and colleagues (Imai et al., 2005; Imai et al., 2008), however, showed that 3-year-olds consistently failed in extending a newly learned verb to a different sort of patient, and this result seemed to be quite robust across languages. The asymmetry may be attributed to task designs that had different levels of cognitive demands, as I conjectured in the literature review section (Section 3.1.1.3). But since these studies vary in many aspects – age, language, choice of linguistic stimuli (e.g. argument type), choice of non-linguistic stimuli (e.g. novel objects or familiar objects), task procedures (e.g. contrast phase), it would require more carefully controlled studies to really pinpoint the factor(s) that contributed to the asymmetry. That said, another possibility, as discussed in Section 3.1.1.2, holds that extension to different agents may be inherently easier than extension to different patients, due to the following observation: the verb tends to imply more about the involvement and nature of the participant realized by the internal argument (i.e. usually the patient), than about that of the participant realized by the external argument (i.e. usually the agent); in this sense, there is a closer semantic relation to the internal argument than to the external argument. This semantic tendency has been theorized by some linguists to reflect some structural distinction, that only internal arguments are arguments of the verb (Kratzer, 1996, 2003; Marantz, 1984). It is possible that this semantic tendency is also expressed in the initial expectations of the young learner such that she tends to think verb meanings include exactly what sort of patient participant, but do not include what sort of agent participant, leading to more reluctance to extend verb meanings to different sorts of patients. Future research is certainly required to test this conjecture.

Second, although extremely challenging, experimental evidence with even younger infants would be necessary to lend more credibility to the argument the current chapter is trying to make – one that argues about the learner's *initial* expectations about verb meanings. We have shown in the current chapter that infants of 22 months are able to extend verb meanings to different participants; but as shown in studies of Chapter 2, infants as young as 18-month-olds are able to learn novel verb meanings guided by the verb-event bias. If we can show even 18-month-olds can succeed in verb extension tasks, this would establish more compelling evidence for our argument. For this purpose, an ongoing experiment in our lab is underway. It is a third experiment of the study presented in Chapter 2, the basic design logic of which is as follows. 18-month-olds are habituated to two pairs of visual-auditory stimuli; in one, a platypus is engaging in a rocking action, paired with a sentence containing a novel verb (e.g. 'it's *dakking*'); in the other, a kiwi is engaging in the same type of action – rocking, paired with a sentence containing a novel noun (e.g. 'it's a *zop*'). When infants reach habituation, the test begins, where some infants receive the same stimuli - *Same Condition*), some receive a re-pairing of visual and auditory stimuli that results in a change in the meaning of the novel noun (e.g. 'it's a *zop*' paired with platypus rocking) – *Noun Switch Condition*, yet some receive a re-pairing that does not result in a change in the meaning of the novel verb (e.g. ‘it’s *dakking*’ paired with kiwi rocking) – *Verb Switch Condition*. Crucially, infants in the Verb Switch Condition would think there is no meaning change *only* if they know that the novel verb *dak* used to label a platypus-rocking event could also be used to label a kiwi-rocking event. The Habituation-Switch Paradigm, by virtue of only requiring visual fixation from the infants’ part, is considered not demanding; compared to the Preferential Looking Paradigm, it may be even less demanding, because it does not require infants to compare and make a ‘choice’ between two competing events. Results from this experiment may be informative for the question at issue.

3.4.4 Some Thoughts on Related Issues

I have been talking about the *initial* representation of verb meanings the learner may entertain. What is worth pointing out, however, is that any learning bias/heuristic can only get learning off the ground, but is not deterministic or exclusive. Just as the correspondence between verb and event is only correlational, verbs that are general in meaning and hence broadly extendible is only part of the story. In fact, many verbs are quite specific, in that they place semantic constraints on the arguments. For example, one can only *eat* things that are edible and *drink* things that are liquid-y; only things with *mind* can think; only horses can *gallop*; *pounding* some cloth sounds weird; *smashing* a napkin is unacceptable; ec cetera. I hypothesized that these finer semantic restrictions are later learned and gradually added to the developing lexicon (Section 3.2.1, part B of my hypotheses). How exactly these restrictions are learned is an equally intriguing question that calls for future research. Here is one possible way: these may be learned by observing that every time a verb occurs, its participant is of a particular type, but not others. For instance, if the learner observes that every time *gallop* occurs it applies to a horse, she may infer that the verb *gallop* not only describes the way its participant is involved, but also specifically describes the type its participant must be, namely, horse sort of things. In other words, it would be a ‘suspicious coincidence’ for the learner if the meaning of *gallop* was more generally applicable but only horses happened to be in the input. The term ‘suspicious coincidence’ (Xu & Tenenbaum, 2007a, 2007b) was used in the Bayesian inference literature to refer to the situation when a random sample of members of a superset (e.g. all kinds of objects) all happen to cluster within one of its subset (e.g. only horses); a Bayesian learner would consider such a situation to be ‘suspicious’ and therefore infer that the subset is the target to be learned (e.g. *gallop* applies only to horses). It has been shown that young learners do use this Bayesian way of inference in learning novel nouns and extending newly learn nouns to other instances (Xu & Tenenbaum, 2007a, 2007b); but its applicability in adjective learning was questioned (Gagliardi, Bennett, Lidz, & Feldman, 2012). Whether or not, and to what extend, this Bayesian inference may be used in learning the finer semantic restrictions of verbs still await further investigation.

Another separate yet interesting question is related to the secondary goal of this chapter, namely, the role of different types of subject on verb learning. Lidz et al. (2009)’s experiments and the current experiment (both using intransitive verbs) demonstrate that the extra information associated with a lexical NP subject may hinder, rather than facilitate, young learners' verb learning and verb extension; and the underlying cause is likely to be the extra processing burden incurred by the extra information. However, for same age (24-month-olds) English-learning infants, Arunachalam and Waxman (2011) showed that lexical NPs (e.g. 'the man is *pilking* a balloon') was better than pronouns (e.g. 'he is *pikling* it') in facilitating verb learning (with transitive verbs); the authors referred to the NP case *richer semantic context* and the pronoun case *sparser semantic context*. These findings may seem contradictory on the surface; however, I think they both speak to the delicate balance between a cue's informativity and possessing load – this balance could be tilted towards different directions under different conditions. For example, in the case of intransitive verbs and one participant event, like Lidz et al. (2009) and the current study, the information provided by the lexical NP subject is to some extent unnecessary, because without that information, it is still clear what is the referent of the subject; In the case of transitive verbs and two participant events, however, unambiguously identifying each referent object may become very important, in which case informativity overrides processing load. What is even more intriguing, Arunachalam, Leddon, Song, Lee, and Waxman (2013) found that learners of argument-drop languages (e.g. Mandarin, Korean) – where noun phrases are frequently dropped when their referents are highly accessible, benefitted more from the sparser semantic context than the richer one (in learning transitive verbs). Recall that such cross-linguistic difference is also seen in Imai and colleagues' studies (Imai et al., 2005; Imai et al., 2008) – that English-learning 5-year-olds succeeded in verb (transitive verbs) extension tasks only when the verb was taught in full-argument structure (pronominal arguments) whereas Japanese-learning 5-year-olds succeeded only when the verb was taught in null-argument structure. In a similar vein, a study of mine, reported in Chapter 5, also suggests that Mandarin-learning toddlers may benefit more from the use of null argument in learning thematic relations. These findings at least suggest that the choice of argument type of the linguistic stimuli plays a very important role, and many factors – grammar, pragmatics, cognitive processing, as well as language-specificity – should be taken into consideration to make the best possible choice.

Chapter 4[25]

A Prelude to the Participant-Argument-Match Bias: Participant Structure Representation in Infancy

As is now well recognized, lexical acquisition cannot be built on pure observations of the word-to-world mapping (cf. Locke, 1690/1959), because the extra-linguistic world does not present itself as observable pieces that are readily mapped onto linguistic units (Landau & Gleitman, 1985). In addition to observational apparatus, therefore, learners must also come into the world prepared with some kind of 'mental zoom lens' that will aid them in slicing the world up and picking out relevant information for further processing. One such 'zoom lens', proposed by (Landau & Gleitman, 1985) in the widely known *Syntactic Bootstrapping Hypothesis (SBH)*, targeting verb learning in particular, is the syntactic environment of the verb. The fundamental idea of this hypothesis is: there are certain principled relations between verb meaning and verb syntax, which is known to the child learner, and thus the learner can use this knowledge to infer verb meaning from verb syntax. In particular, it is proposed that the learner expects the argument structure of a sentence and the participant structure of its event to align – first, the number of participants and arguments are expected

[25] The content in Chapter 4 was also reported in a conference proceeding paper: Wellwood, A., A. X. He, J. Lidz, & A. Williams (to appear). Participant structure in event perception: towards the acquisition of implicitly 3-place predicates. Proceedings of the 38th Annual Penn Linguistics Colloquium (PLC), 21(1). Philadelphia, PA: University of Pennsylvania.

to match up and second, the thematic roles of participants are expected to match up with the syntactic position of arguments. The second aspect will be discussed in Chapter 5. The current chapter focuses on the first aspect, and will refer to it as the *Participant-Argument-Match bias*, or *PAM*.

When acquiring a semantics, the learner relates her experience of the speech (i.e. linguistic representation) to her experience of the world (i.e. extralinguistic representation). In learning verb meanings, according to the PAM hypothesis, the learner expects to interpret a novel verb with n argument noun phrases (linguistic representation) as describing an event with m (m = n) participants (extralinguistic representation). When this hypothesis is studied in the laboratory, the latter part of this relation is more often presumed than independently tested. But what are the 'participants' in an event? The world itself is not labeled with participant roles; these are aspects of the concept under which people view a stretch of the world, and the same thing can be viewed in many different ways. Therefore, the hypothesis that the learner uses PAM as a guide in learning novel verbs is only falsifiable to the extent that we are confident about how the learner views the world, that is, how the learner represents the extralinguistic input.

This chapter investigates the learner's extralinguistic representation, as a necessary prelude to testing the PAM hypothesis. The PAM hypothesis says that participants match arguments for any number, but most evidence for PAM is limited to cases of 1 and 2 – the learner expects a 1-argument sentence to name a 1-participant event and a 2-arugment sentence to name a 2-participant event. But on these cases, there are other possible learning heuristics that happen to agree with PAM; for instance, heuristics like ‘transitive verbs name causative events, intransitive verbs name non-causative events’. To distinguish PAM from such heuristics, we should look at other cases, like the case of 3, where only PAM will predict a match. This chapter examines such a case: 2-argument (n = 2) sentences that plausibly express 3-participant (m = 3) concepts. For example, ‘Anne stole the toy’ expresses a THEFT concept, which entails a *victim*, alongside the *thief* and his *loot*, and it seems very plausible to consider this *victim* a ‘participant’ in the event when we view it as s stealing, instead of as a picking-up. Such cases present a potential mismatch between the number of arguments (n = 2) and the number of participants (m = 3), posing a potential challenge to the PAM hypothesis. Investigations of such cases will extend previous research on PAM beyond cases of 1 and 2, and shed light on the effectiveness of this bias in learning verbs like *steal* that are implicitly three-place.

In what follows, I will begin with a review of previous research that lend support for the PAM hypothesis, discussing important limitations in them that motivated the current investigation on event representation (Section 4.1); I will then present in detail the types of events that the current study is interested in (Section 4.2); following that, I will discuss our attempts in developing methods to experimentally examine people’s event representation (Section 4.3); subsequently I will present four experiments with a series of different event types testing adult subjects, as some preparatory studies to select the best event types to test infants (Section 4.4 to Section 4.5), and then three experiments testing infants’ extralinguistic representation, following the lead of (Gordon, 2003) (Section 4.6 to Section 4.7); finally, I will conclude this chapter with a general discussion in Section 4.8.

4.1 Background

This chapter aims to develop an adequate test of the participant-argument-match hypothesis (PAM) of verb learning, by examining the learner's extralinguistic representation. With this goal, this section first reviews existing experimental evidence for this hypothesis (Section 4.1.1), then discusses important limitations in the literature that motivated the current investigation on representations of plausibly three-participant events (Section 4.1.2).

4.1.1 Literature: Evidence for PAM

The participant-argument-match hypothesis (PAM) has received important initial evidence (Fisher, 1996; Fisher et al., 2010; Fisher, Hall, Rakowitz, & Gleitman, 1994; Gleitman, 1990; Lee & Naigles, 2008; Lidz et al., 2003; Naigles, 1990; Naigles & Kako, 1993; Yuan & Fisher, 2009; inter alia). Here I review some of them that provided compelling evidence for the role of the number equation in guiding novel verb learning.

A classic study by (Naigles, 1990) gave the first direct evidence for PAM. In a preferential looking task, 25-month-olds were familiarized with a complex event in which a duck was pushing a bunny down while both of them were swinging their arms around; they also heard a concurrent sentence description of the scene - half of them heard a transitive sentence 'the duck is *gorp*ing the bunny', and the other half heard an intransitive sentence 'the duck and the bunny are *gorp*ing'. At test, these children were shown on opposite sides of the screen two sub-events decoupled from the complex one – a duck-pushing-bunny-down event, and a duck-and-bunny-swinging-arms event; and they were asked to look at either side of screen by the prompt sentence 'where's *gorp*ing'. Results were: those who heard the transitive sentence looked longer at the sub-event in which the duck pushed the bunny down than at the sub-event in which the duck and bunny were swinging their arms, while those who heard the intransitive sentence showed the opposite looking preference. These results suggested that children around 2 years of age were able to use the number of argument noun phrases in a sentence as cues to make inferences about the meaning of a novel verb: those who heard the novel verb in a transitive context conjectured that *gorp* meant something like "force to squat", whereas those who heard it in an intransitive context conjectured a meaning like "swing the arms". These results further suggested that children were building some syntactic representations of the linguistic stimuli (i.e. a verb and its argument structure), rather than simply counting nouns, because both the transitive and intransitive sentences contained two nouns, but only the transitive sentence had two arguments.

Yuan and Fisher (2009) pushed this finding one step further by showing that sentence structures alone carry information about meaning, and learning of novel verb meanings could occur in absence of concurrent scenes that presumably hinted at the verb's semantic content. In a preferential looking task, 28-month-olds were played some dialogue between two females, containing eight sentences with a novel verb (e.g. *blick*) – half of the children heard a dialogue with transitive sentences (e.g. 'Jim is gonna *blick* the cat', 'Mary was *blick*ing the man'), and the other half heard a dialogue with intransitive sentences (e.g. 'Jim is gonna *blick*', 'Mary was *blick*ing'); importantly, they received the sentences via watching two female characters talking to each other on the screen, but saw no presentation of the scenarios described by the sentences. At test, children were then shown two different events on opposite sides of the screen – an event where one person pulls along another person's leg who was seated in a chair (i.e. a two-participant event), and an event where a person was swinging her arm (i.e. a one-participant event); and they were asked to look at either side of the screen with the prompt 'where's *blick*ing'. Results were: those who heard the transitive dialogue looked reliably longer at the two-participant event than the one-participant event, whereas those who heard the intransitive dialogue did not show a looking preference between the two events (see more discussion about this result in Section 4.1.2). These results suggested that toddlers were able to glean some structural information from sentences alone and later use this information to guide their attention to certain candidate event. A second experiment of this work showed that learning was successful even when there was a delay (1 to 2 days) between the dialogue phase and the test phase, further supporting the robustness of this structurally guided learning.

Yuan et al. (2012) demonstrated this ability to use the number of syntactic arguments to infer the meaning of a novel verb in children under the age of two. In a preferential looking task, 19- and 21-month-olds saw two events played on opposite sides of the screen – one event where a boy caused another boy's to back bend by pushing him (i.e. a two-participant event), and another event where a boy performed jumping-jack motions (i.e. a one-participant event). Prior to seeing these events, children heard a sentence describing the upcoming scene in future tense (blank-screen displayed), and concurrent with these events, they heard a sentence describing the scene in present tense; half of the children heard a transitive sentence (e.g. 'he's gonna *gorp* him', and 'he's *gorp*ing him'), and the other half heard an intransitive sentence (e.g. 'he's gonna *gorp*', and 'he's *gorp*ing'). At test, children were asked to look at either side of the screen with a prompt sentence 'find *gorp*ing'. Results were: those children who heard a transitive sentence looked more to the two-participant event than to the one-participant event, but those who heard an intransitive sentence did not show a preference (see more discussion about this result in Section 4.1.2); same results held for both 19- and 21-month-olds. Notice that in this study, arguments were denoted by pronouns (i.e. *he* and *him*) whose referents were ambiguous; this manipulation, removing the semantic content otherwise provided by lexical noun phrases, was to isolate the number of arguments as an independent cue to verb meaning (also see Fisher (1996) for a similar manipulation). Young toddlers' success in this task, therefore, provided strong evidence for the guiding role of argument number.

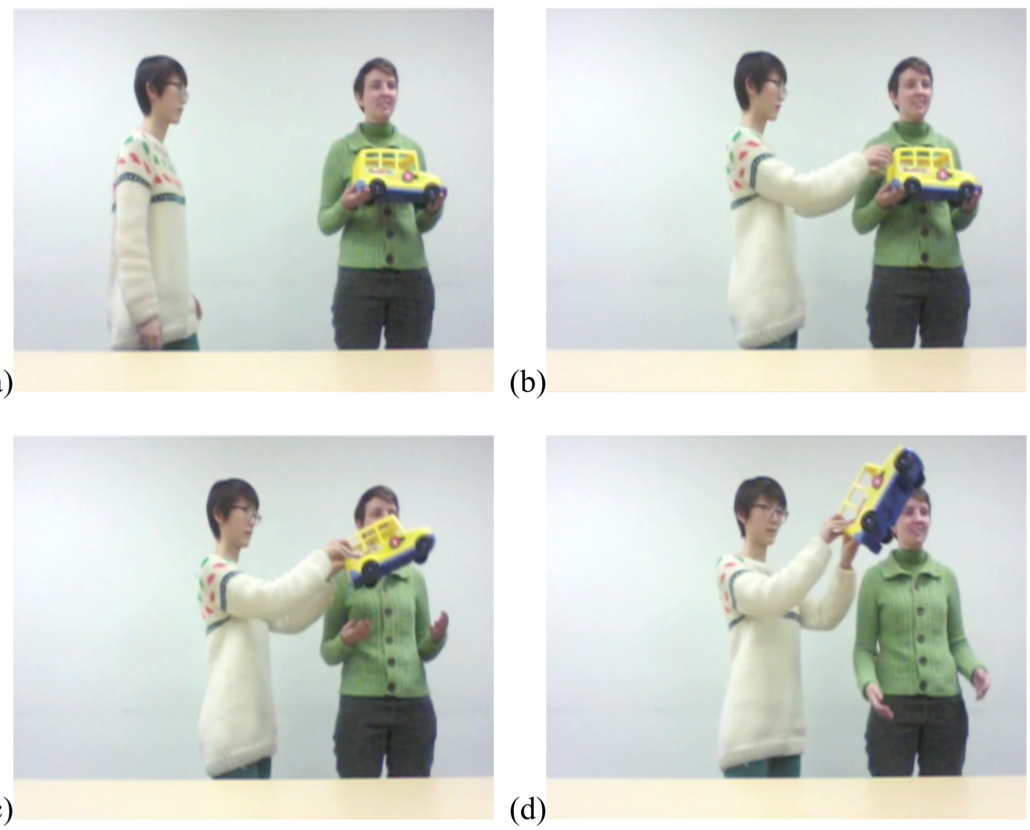
Importantly, Yuan et al. (2012) raised a possible alternative interpretation for children's performance in the above-reviewed experiment: those children might not encode the nonlinguistic stimuli into some event structure with participants, but simply count the number of entities in each scene, and made their choice based on a number match between arguments and visible entities, rather than a number match between arguments and participant roles. To evaluate this alternative, the authors conducted a second experiment: they added a bystander to the one-participant event such that this event also had two people, just like the two-participant event; but the second person, standing idly by, is not a participant of the event. Results were the same: both 19- and 21-month-olds significantly preferred the two-people two-participant event over the two-people one-participant event for the transitive clause. This 'bystander' manipulation highlighted the point that not every entity in a scene is necessarily as a participant role, and suggested that a) young children were actively building some representation of the extralinguistic world, encoding some entities as participants, some not, and b) the PAM hypothesis operates on the notion *participant* as it is mentally represented. I will discuss what it's meant by the notion *participant* in Section 4.2.1.

4.1.2 Motivations for the Current Study

Previous studies, although providing important initial evidence for the PAM hypothesis, are limited in two ways.

First, previous investigations are limited to cases where the number is 1 and 2. On cases of 1 and 2, there are many possible hypotheses that would agree with PAM; for instance, some heuristic that says transitive verbs name causative events would agree with PAM's prediction on cases of 2. What is distinctive about Pam, is that is applies equally for all numbers. It is just about number match, and nothing more specific. Therefore, if we want evidence for PAM in particular, evidence that distinguishes PAM from other conjectures that partly overlap with it, we should look at the case of 3. What is more worrisome, within studies that tested 1 and 2 cases, only results with the 2 case have been consistent and robust: children have been shown to prefer a two-participant event when hearing a two-argument sentence (e.g. Naigles, 1990; Yuan & Fisher, 2009; Yuan et al., 2012). Results with the 1 case, however, have been quite inconsistent: when hearing a one-argument sentence, children sometimes did not show preference of a one-participant event over a two-participant event; for example, in both Yuan and Fisher (2009) and Yuan et al. (2012) reviewed in the previous section, children who heard the intransitive sentence attended to the two-participant event for about the same time as to the one-participant event. The authors attributed this finding to multiply possible ways to perceive a two-participant event: the two-participant event presented in these experiments could in principle be viewed as a single two-participant event (e.g. A spins B), or as one containing a one-participant sub-event (e.g. B spins) (cf. Bunger, 2006; Fernandes, Marcus, Di Nubila, & Vouloumanos, 2006; Fisher, 2002; Naigles & Kako, 1993); therefore, children could find a one-participant event in both scenes, resulting in equal amount of attention. This observation points out the importance of independently studying how a certain given scenario is viewed, which is the next point I am going to discuss.

A second and perhaps more important limitation of previous research on PAM is the lack of independent test on how the learner represents the nonlinguistic stimuli used in the experiments. PAM hypothesizes that the learner uses the number of arguments in the sentence to infer the number of participants in the event, which is essentially about establishing a relation between some linguistic representation of the sentence and some extralinguistic representation of the world. The extralinguistic part of this relation, however, is usually presumed in previous research: for example, a scenario where a person drags the leg of another person who is seated in a chair is presumed to be viewed as a two-participant event. Yuan et al. (2012)'s 'bystander' manipulation is among the very few that takes into consideration the learner's representation of the extralinguistic world; even in their experiment, however, that the bystander is encoded as a non-participant is still taken for granted, and is used to test the PAM hypothesis: it was presumed the event with the bystander is a one-participant event, and the other event a two-participant event; if the child looked more to the presumed two-participant event than the presumed one-participant event when hearing a transitive sentence, then, this was taken as evidence for PAM. Such presumptions are fair, based on intuitions about how such events are represented; but they still need to be justified. After all, any given stretch of the world, however simple it is, can be viewed under many concepts that differ in the number of participant roles; without independent experimental evidence on how a certain given scene is viewed, our intuitions stay merely as intuitions. For example, the simple scene in Figure 4.1, can be viewed as a standing that has one participant (the stander), a holding that has two participants (the holder, the held), a picking-up that has two participants (the picker, the picked), a stealing that has three participants (the thief, the loot, the victim), etc. Therefore, before using simple scenes like this in experiments testing for PAM, we need to be confident about how such scenes are viewed by the child learner.



(a) (b) (c) (d)

Figure 4.1: Snapshots of a stealing scene – (a) - (d) in temporal order

To sum up, the PAM hypothesis states that the argument noun phrases in a sentence exactly match in number with the participants in the event. But the extralinguistic world itself is not labeled with participants. Participant roles are in people's mental representations: they are aspects of the concept under which people view a stretch of the world. (See Section 4.2.1 for a discussion on what is a participant.) Crucially, the same piece of world can be viewed in many different ways. Thus the hypothesis that children use PAM is only testable if we are confident about how the child learner views events. The current study, therefore, aims to investigate the learner's extralingusitic representation – how many participant roles are encoded for a certain given scene, as an important prelude towards an adequate test for PAM. To move beyond cases of 1 or 2, this current investigation looks at cases where a scene is plausibly viewed under a 3-particpant concept (m =3, plausibly), for example, givings, stealings, jimmyings, and beanings. Among these, some are naturally described by sentences that have the same number of arguments as plausible participants (e.g. givings) (n = m = 3), whereas some are naturally described by sentences that have fewer number of arguments than plausible participants (e.g. stealings, jimmyings, and beanings) (m = 3, n = 2), creating a potential number mismatch between arguments and participants, thus posing a potential challenge for PAM. I will discuss these cases in more detail in the following section.

4.2 Cases of Interest

I will begin this section with a discussion on how we characterize a 'participant', then move onto the specific cases this study will investigate.

4.2.1 Defining 'Participant'[26]

To characterize what we mean by a *participant*, take the event of giving for instance. Every event of giving occurs at some time, in relation to various locations and individuals, involving transfer of some property and some speed and direction of the transfer, etc. These relations are therefore *entailed* by the event concept GIVE, which is expressed by the verb *give*; see the outer circle of Figure 4.2. Among the entailed roles, some are special, in the sense that they are psychologically distinguished or foregrounded in some way – a privileged subset. We suppose that for *give* this privileged subset includes just the agent, the theme and the recipient; that is, the giver, the given, and the given-to; see the inner circle of Figure 4.2. When we represent an event with the concept GIVE, these roles are *explicit* - we might think of them as the *explicit constituents* of a structured event representation associated with the verb. This is not to deny that all givings have locations, times, or transfer of property, for example; these roles are entailed by the concept GIVE, but are not separately represented. What exactly it means to be an *explicit constituent* of a psychological representation is not clear, but something like this notion is presupposed in much grammatical theorizing. A

[26] This is also discussed in Williams (2015).

perhaps more intuitively analogy is this: when we take “☺” as a representation of a normal human face, only the mouth, the eyes and the outline are explicitly represented – explicit constituents; we can mean this as a representation of a face without denying that faces have noses.

(4-1) ***Participant roles*** are a privileged subset of the roles entailed by a concept.

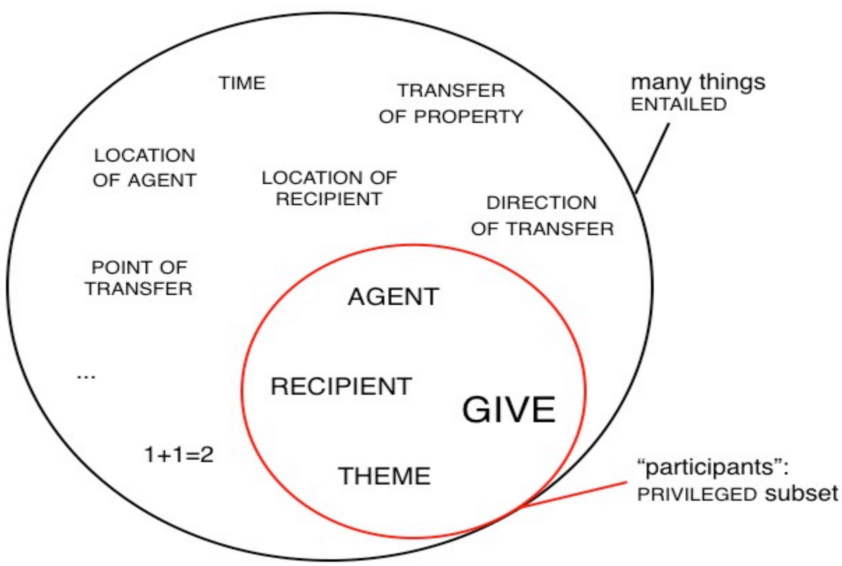


Figure 4.2: Entailed roles and participant roles of the concept GIVE

4.2.2 Plausibly 3-Participant Events

In previous research testing PAM for the n = 1 or 2 cases, it usually presents two candidate scenes that presumably differ in the number of participants – a one-participant event versus a two-participant event, and teach children a novel verb either in a transitive sentence or in an intransitive sentence; If children's choice of event is modulated by the sentence structure they receive, it is considered to support PAM. Consider now that we are to develop a test for PAM for the n = 2 or 3 cases. I will begin with a simper case (Section 4.2.2.1), and then move on to a trickier case (Section 4.2.2.2).

4.2.2.1 Plausibly 3-Participants, 3-Arguments

Suppose the two candidate scenes are a giving scene (Figure 4.3) and a hugging scene (Figure 4.4). Following the 'bystander' manipulation of Yuan et al. (2012), to control for counting the number of entities rather than participants (see relevant discussion in Section 4.1.1), both scenes have three salient entities – Anne, Betty, and a teddy bear; but not all of them are necessarily participants. By our intuition, the giving scene in Figure 4.3 is plausibly viewed with three participants – the giver, the given-to and the given; and the hugging scene in Figure 4.4 is plausibly viewed with two participants – the hugger and the huggee, with the teddy bear as a 'bystander'. In an experiment testing PAM, following previous research, we present these two scenes on opposite side of the screen, and divide children into two groups, one receiving a transitive sentence 'Anne *gleeb*ed Betty' and the other a ditransitive sentence 'Anne *gleeb*ed Betty a teddy'. Predictions are: if children are guided by PAM (**condition a**), *and* if they indeed view the two scenes in the ways we think are plausible (**condition b**), then, those who heard the transitive sentence will look more towards the hugging scene (presumably two-participant), whereas those who heard the ditransitive sentence will look more toward the giving scene (presumably three-participant).

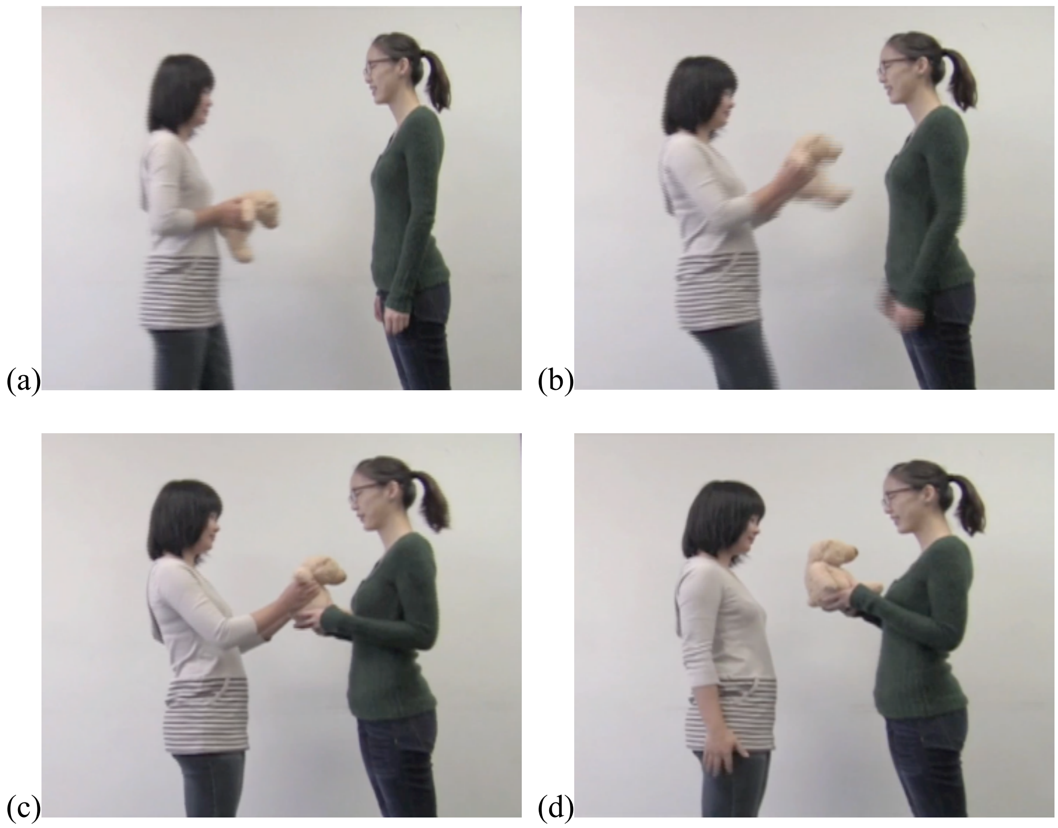


Figure 4.3: Snapshots of a giving scene – (a) - (d) in temporal order

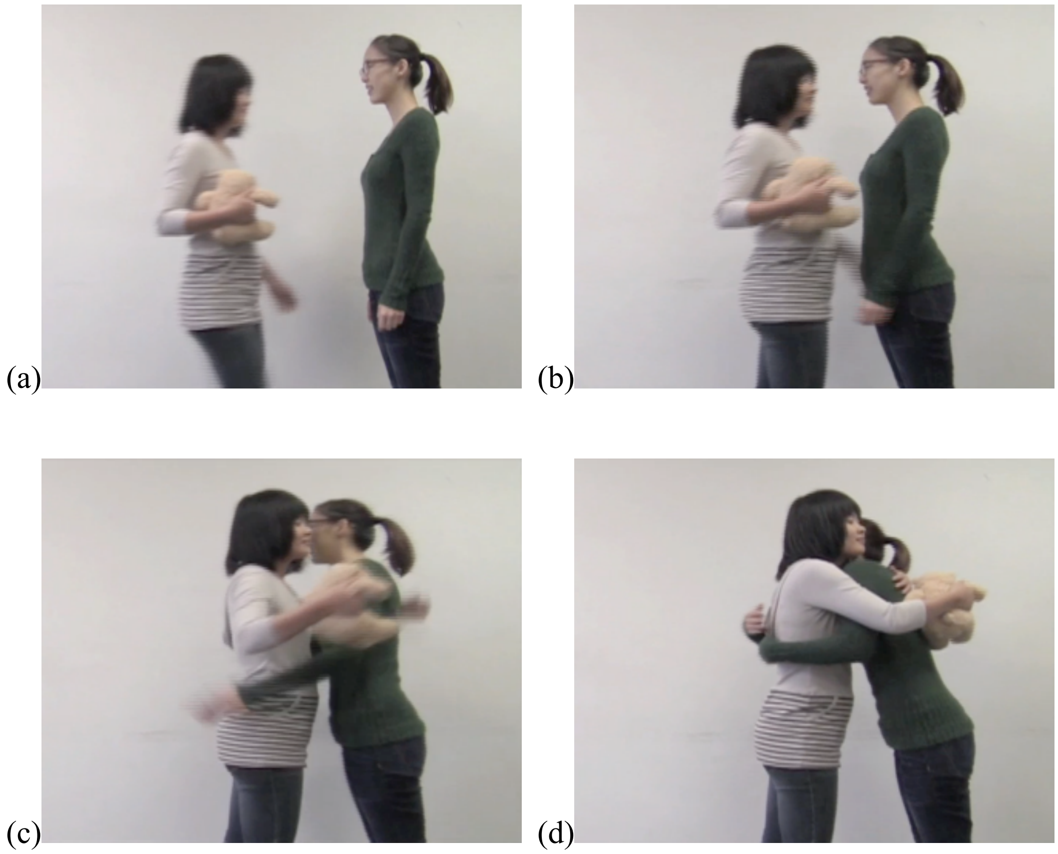


Figure 4.4: Snapshots of a hugging scene – (a) - (d) in temporal order

Crucially, if the experiment is to test **condition b**, we must make sure **condition a** holds. This is not trivial, because it is not guaranteed that these scenes are viewed in the above-mentioned ways: for instance, the giving scene in Figure 4.2 could be viewed under a two-participant concept GIVE', one that is satisfied only by givings, but leaves the entailed role of the thing-given implicit, roughly as the nose is implicit in our drawing of a face; similarly, it is not impossible to view the hugging scene in Figure 4.3 under a three-participant concept HUG', one that additionally encodes the teddy as a participant. Independent experimentation examining the learner's representations of these scenes, therefore, are necessary, for our purpose of developing a test for PAM.

Now consider how verbs like *give* and *hug* may be acquired under the guidance of PAM. Giving scenes can be naturally described with sentences containing three arguments, like (4-2), and hugging scenes can be naturally described with sentences containing two arguments, like (4-3).

(4-2) Anne gave Betty a teddy.

(4-3) Anne hugged Betty.

Suppose the learner indeed views these scenes in the ways that we think they do, then, she will be able to learn the target meanings of *give* and *hug* from input sentences like (4-2) and (4-3), under the guidance of PAM: specifically, guided by PAM, sentence (4-2) leads the learner to hypothesize a three-participant concept for the yet-unknown verb *give*, and she can find such a concept GIVE in her encoding of the extralinguistic world; similarly, sentence (4-3) leads her to hypothesize a two-participant concept for the yet-unfamiliar verb hug, and she can find such a concept HUG in her encoding of the extralinguistic world. See Figure 4.5 and Figure 4.6 for illustrations.

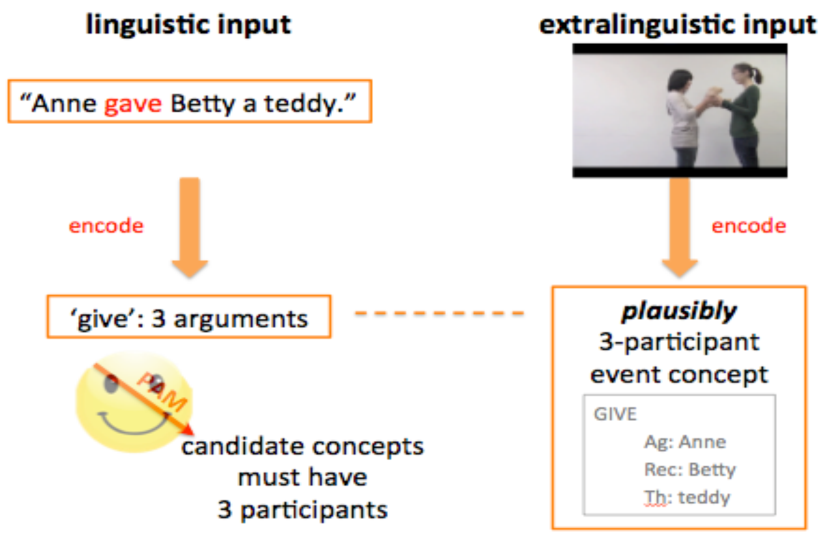


Figure 4.5: Learning the verb *give* under the guidance of PAM

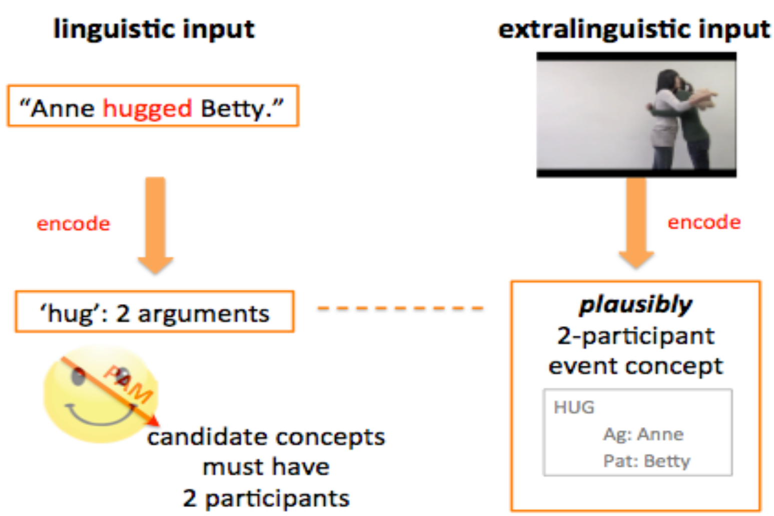


Figure 4.6: Learning the verb *hug* under the guidance of PAM

Suppose, instead, the child learner views these scenes in some alternative ways (e.g. givings under a two-participant concept GIVE’, and huggings under a three-participant concept HUG’), rather than in the ways we think they do; or, to back up a little, suppose the learner *still* views these scenes in the ways we think they do, but does not overwhelmingly do so, instead, is equally likely to view them in GIVE and GIVE’, and HUG and HUG’. Then, PAM will not be an effective learning guide to narrow down possible meanings of the verbs, because when the learner hears (4-3), she hypothesizes a space of two-participant concepts for the yet-unknown verb *hug*, which does not exclude the two-participant concept GIVE or include the three-participant concept HUG’. Therefore, to make any argument for PAM, we must be confident about which concept the learner is more likely to view a certain given scene under.

4.2.2.2 Plausibly 3-Participants, 2-Arguments

Scenes like givings and huggings, despite the need for independent test on event representations, may be less of a concern for learning guided by PAM, because the plausible ways of viewing the scenes (according to our intuition) have participant structures that match in number with the argument structures of sentences describing them: givings, plausibly 3 participants, and 3 arguments (4-2); huggings, plausibly 2 participants, and 2 arguments (4-3). Another case, however, is more concerning: transitive sentences (4-4, 4-5, 4-6) whose events, in our judgment, seem plausibly to be viewed under concepts with more than two participant roles (Fillmore, 1982; Goldberg, 1995).

(4-4) Anne *stole* a toy.

(4-5) Anne *jimmied* the box.

(4-6) Anne *beaned* Betty.

The above sentences entail a third role that corresponds to no overt argument. Every stealing (see Figure 4.1 in Section 4.1.2) has a victim, in addition to a thief and some loot; every jimmying (see Figure 4.7) has a lever, in addition to its wielder and what the lever is used to open; and every beaning (see Figure 4.8) has a projectile, in addition to its thrower and the person whose head is hit with it. It also seems plausible that children will view such events under a concept that foregrounds this third role as a participant role, alongside the two others. Plausibly, they are less likely to view them under two-participant concepts that entail the third role, but leave it at that, as for the entailed roles of time or place.

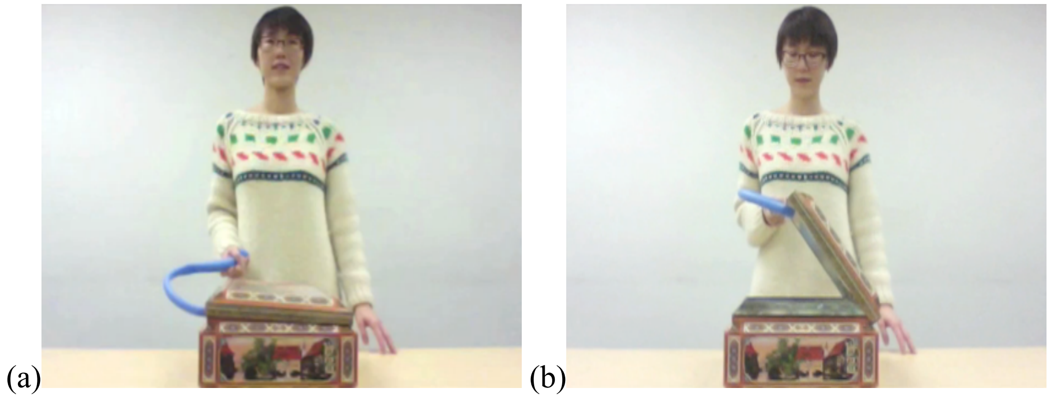


Figure 4.7: Snapshots of a jimmying scene – (a) - (b) in temporal order

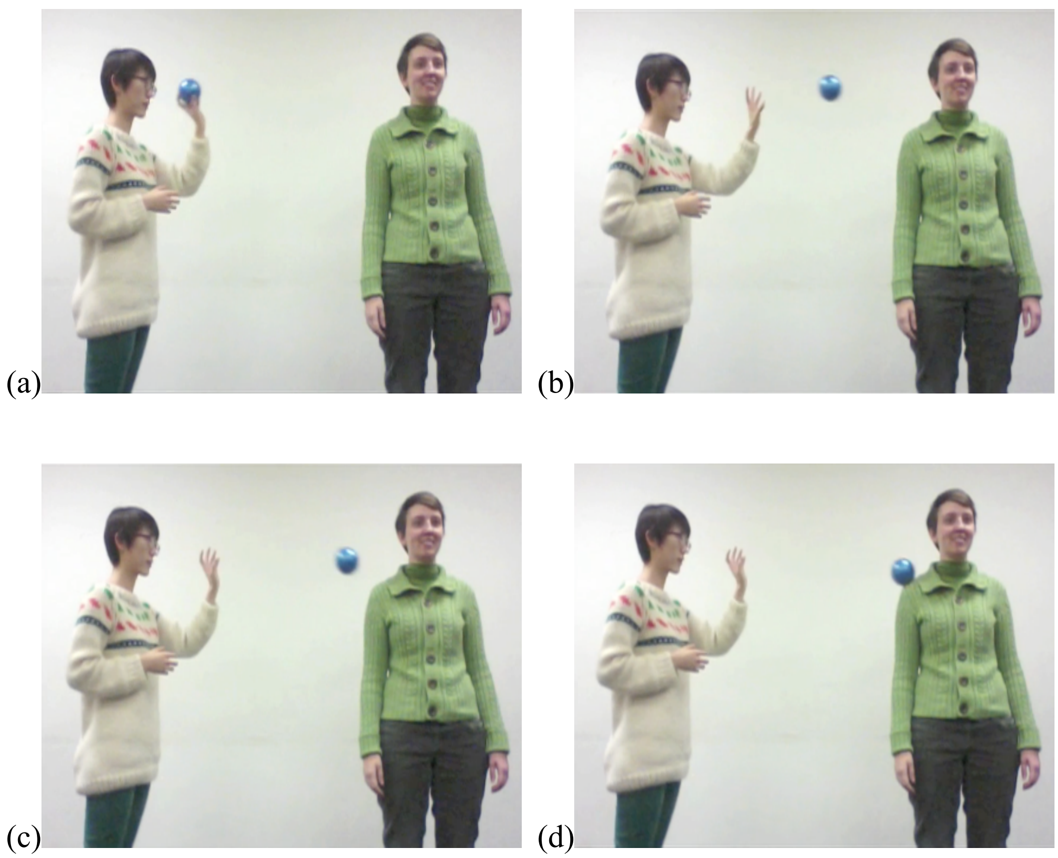


Figure 4.8: Snapshots of a beaning scene – (a) - (d) in temporal order

If this is correct – that these scenes are readily viewed by the learner under three-participant concepts, then PAM makes a prediction, that the learner will fail to understand sentences like (4-4), (4-5), or (4-6) and the verbs in those sentences: guided by PAM, these two-argument sentences will lead the learner to look for event concepts that have two participant roles, excluding any three-participant concepts from her hypothesis space of verb meanings – that is, the learner will *not* consider the three-participant concepts that she views the scenes under (STEAL, JIMMY, BEAN) as candidate meanings for the verbs (*steal*, *jimmy*, *bean*). What if the learner readily shifts her view of the scenes from three-participant concepts to two-participant ones? For example, instead of viewing stealings under the concept STEAL, she may view it under the concept PICK UP, which isolates only two aspects of the event as participant roles. In such a case, under the guidance of PAM, the learner will interpret the verbs as expressing more general concepts that do not entail the third role, such as PICK UP instead of STEAL, OPEN instead of JIMMY, or HIT instead of BEAN. And this misunderstanding will persist until the counsel of PAM is abandoned. Thus we have a clear test of PAM, but only if we can first show that children view thefts, jimmyings, and beanings as we expect, under a three-participant concept. See Figure 4.9 for an illustration of the learning problem.

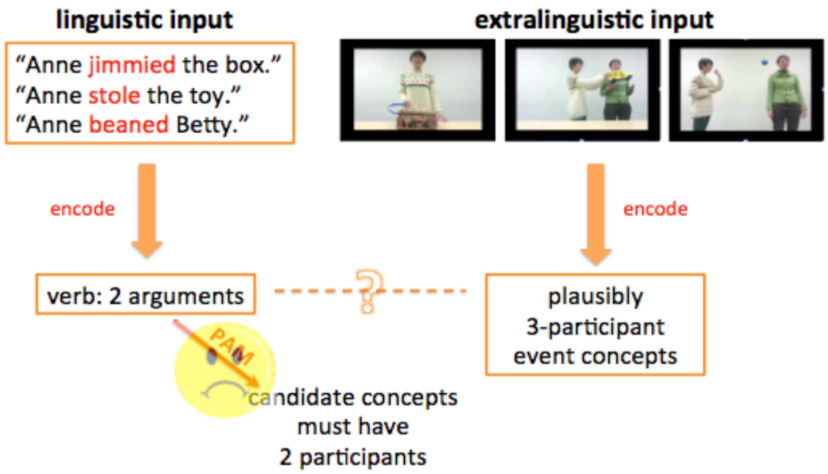


Figure 4.9: Trouble in learning *jimmy*, *steal* and *bean* under the guidance of PAM

To sum up, this study will investigate event representations of these plausibly three-participant events, for both the number-match case (e.g. givings, with huggings as control) and the number-mismatch cases (e.g. stealings, jimmyings, beanings). See the next section for our attempts towards some appropriate methods of testing event representation.

4.3 Methodological Attempts

At present we have no sure test for whether someone is viewing an event under a three-participant concept, versus a two-participant concept. But we can take steps to increase our confidence.

Suppose a subject views two scenes that differ minimally. One exhibits an event *e* to which an individual *x* bears some relation *R* – for example, an event of opening a box (*e*), in which a lever (*x*) is used as an instrument (*R*); see Figure 4.10(a). The other scene is the same, except that *x* is inert – the lever is present but unused; see Figure 4.10(b). Then if, according to some behavioral measure which we take to be informative, the subject does *not* respond to this difference as important, we may infer that she did *not* view the first scene under a concept which has *R* as a participant role – in our example, she viewed the first scene as an opening, and not more specifically as a jimmying (that is, as an opening-with-lever). Conversely, if she *does* respond to the difference, it remains possible that *R* is a participant role in her view – in our example, she viewed the first scene as a jimmying that specifically has the lever as a participant. To make sure that it is difference in *participant structure* that drives the learner response, not some general perceptual difference, we also need to test her on another pair of minimally different scenes, like jimmyings from different orientations (Figure 4.10(a) and 4.10(c)).

(a) 

(b)  (c) 

Figure 4.10: Illustration of two types of contrasts – opening vs. jimmying & opening-from-left vs. opening-from-right – (a) opening from left, lever active; (b) opening from left, lever inert; (c) opening from right, lever active. These are **Stimuli of Experiment 2** (Section 4.5)

Gordon (2003) makes a similar argument about event representation in pre-linguistic infants, presenting pairs of minimally different scenes with the Habituation-Switch paradigm (Casasola & Cohen, 2000; Werker et al., 1998; Younger & Cohen, 1986), one during the habituation phase and one during the test phase; in this method, infants' decreased attention from habituation to test is taken as no response to the difference, and their recovery of attention is taken as response. Specifically, in the habituation phase, Gordon showed infants a series of videos in which one woman (Anne) hands another (Betty) a teddy bear, or in which Anne hugs Betty while holding a teddy bear. When infants reached the habituation criterion, they were presented with a new video. In the Same condition, this was a token of the same type of scene (that is, Anne hands Betty a teddy bear, or Anne hugs Betty), but a token that they wasn't seen before; in the Switch condition, this was a token of a different type of scene, in which the same basic action is performed, but without the teddy bear. Of interest was whether infants' attention was recaptured in response to the new videos (i.e. dishabituation). Results were: infants dishabituated – that is, they attended more towards the new stimuli than they had at the end of the habituation phase – in the Switch condition for the giving but not for the hugging.

Gordon interpreted these results as suggesting that infants thought something interestingly different was happening when the teddy bear was removed from givings but not when it was removed from huggings; and this in turn suggested that they represented the hugging scenes under a concept that does not entail the role of the teddy, and the giving scenes under a concept that does. Had the infants failed to notice or care about the disappearance of the teddy bear from the scenes of giving, it would seem unlikely that they were viewing that scene under a concept that entails that role.

Gordon's method shows, we believe, that the noticed role is entailed by the child's view of the event, while the unnoticed role is not. But certainly it cannot prove the stronger conclusion, that the noticed role is furthermore a participant role (i.e. an entailed role that is privileged). Nonetheless, it does lend some credence to that conclusion if we presume, as seems reasonable, that one will not attend equally to every role entailed by the concept under which one views a thing. Human faces have noses, but we attend more to the eyes, which seem to have a distinguished role. Thefts in general have locations, but perhaps when we view an event as a theft, we are more likely to attend to the thief and the victim. Thus in our own work with infants, reported in detail in Section 4.6 and Section 4.7, we adapt Gordon's methods to our own purposes, in designing tests of PAM.

Before turning to young infants, we first conducted some preparatory studies, in which we sought to find a correlative measure of the participant-structure distinction in adults. Following the same design logic – that is, presenting subjects some pairs of minimally different scenes that either differ in participant structure or merely differ in some perceptual feature (e.g. orientation of action) and testing to which difference subjects respond, we used a similarity judgment task to assess the adult view of the stimuli we prepared for studies with infants. I will present a series of such experiments with adults in Section 4.4 and Section 4.5.

4.4 Experiment 1: Adults, Number-Match

In this section and the next, I will present experiments with adults on the following event types that are plausibly viewed under three-participant concepts: first, the number-match cases, givings, with huggings as control, (Experiment 1, Section 4.4); and then, in the next section (Section 4.5), I present three experiments with the number-mismatch cases, jimmyings (Experiment 2), stealings (Experiment 3), and beanings (Experiment 4).

In Experiment 1, we had adults perform a similarity judgment task with stimuli like those that Gordon (2003) used for infants – givings with/without teddy and huggings with/without teddy. The goal is to see how adults view the teddy bear in giving type of events and hugging type of events: whether the teddy bear is psychologically privileged in adults' view of these events such that a difference in its existence elicits some response.

4.4.1 Participants

Twelve adult participants were recruited from the University of Maryland undergraduate community. Participants were invited to our lab on campus, where they were asked for consent in line with the University of Maryland's IRB protocols. They were debriefed about the goals of the study afterwards. Participants received course credit or $10 for their participation.

4.4.2 Stimuli

The visual stimuli were pre-recorded events involving two girls and a teddy bear. There were four types of events: a) giving-with-teddy, where Anne held a teddy in her hands and handed it over to Betty (Figure 4.11(a)); b) giving-without-teddy, same as a) except the teddy was not included (Figure 4.11(b)); c) hugging-with-teddy, where Anne hugged Betty with a teddy in her hand (Figure 4.11(c)); and d) hugging-without-teddy, same as b) except the teddy was not included (Figure 4.11(d)). Two different tokens were recorded for each event type, by having the actors performing the events two times, allowing naturally occurred differences (e.g. slightly different postures, positions of hands, facial expressions, etc.). Thus, there were 8 tokens in total, each lasting about 6 seconds. There was no concurrent audio; in other words, the stimuli were presented in complete silence.

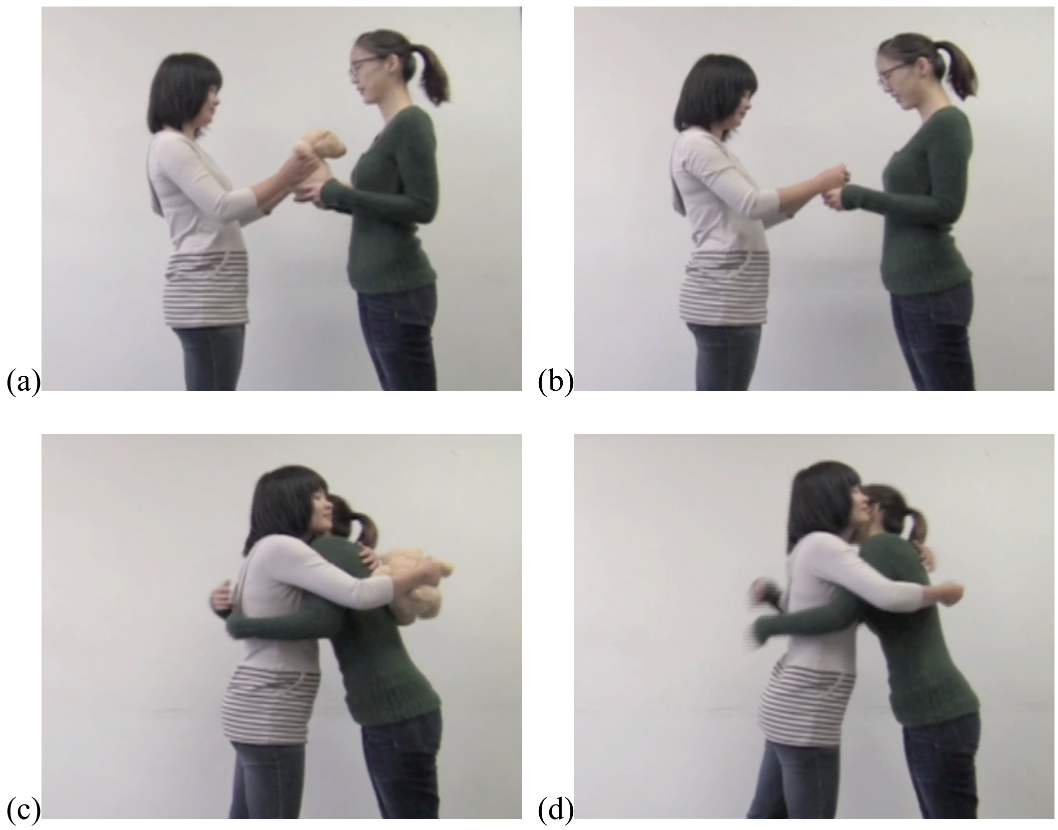


Figure 4.11: Stimuli for Experiment 1

4.4.3 Apparatus

The experiment was run in the Psyscope program (Build 57) with a Mac OS desktop.

4.4.4 Design

This experiment used a Similarity Judgment Task (SJT), in which adult subjects were asked to judge the similarity between two videos. The structure of a single trial involved showing two videos one after another on a computer screen. Immediately following presentation of the videos, the question 'How similar were these two videos' was presented on the screen, together with a 1–7 scale where '1' was marked as 'totally not similar' and 7 as 'totally similar'. Figure 4.12 illustrates the sequence for one trial. Participants were instructed to make judgments as quickly as possible, without sacrificing accuracy; if they were unsure, they had the option to replay the video by pressing the spacebar.

There were 32 trials in total, among which 8 trials were *giving-type comparisons* – giving-with- vs. without-teddy, 8 were *hugging-type comparisons* – hugging-with vs. without-teddy, and the remaining 16 trials were *token comparison* – different video tokens for each scene type (e.g. give-with-teddy token 1 vs. give-with-teddy token 2). These three comparisons were within-subject conditions. Each participant completed all 32 trials in pseudo-randomized order.

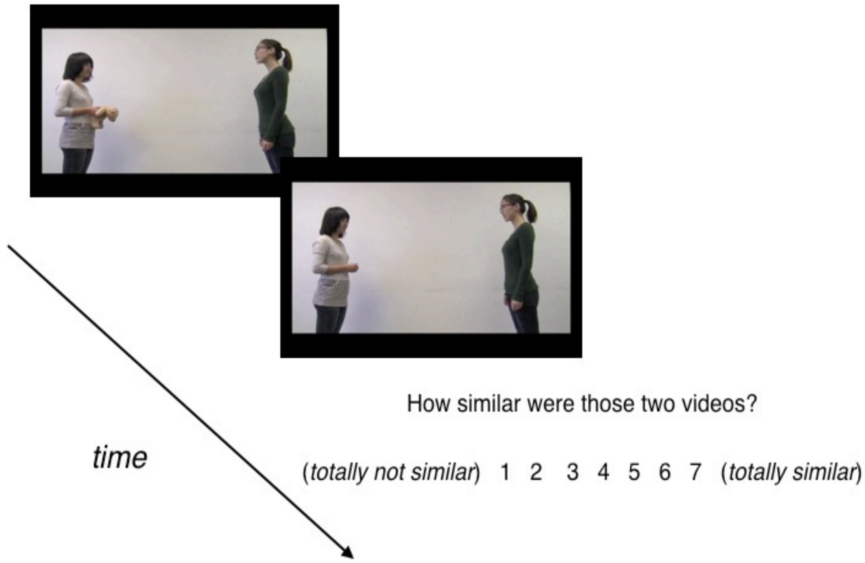


Figure 4.12: Schematic of the trial structure in Experiment 2

4.4.5 Measurements

Participants' rating score was our explicit measure. We also measured the time taken to make a rating decision, as an implicit measure of the processes involved in making the decision. In general, we suspected that the criteria used to make a similarity judgment are not likely to be the same as are used to encode a scene. Explicit judgments of similarity may reply on any number of different criteria – certainly not just those implicit in our initial perception of the scene. Rating scores were measured for all trials, but reaction times (RTs) were only measured on trials without replay. The rationale for this is that reaction times after seeing the videos more than once may not be an accurate measure of the same sort of processing as seeing the videos once. In this sample of 12 participants (32*12=384 trials), the replay option was used on 51 trials in total; those trials were excluded from the RT analysis.

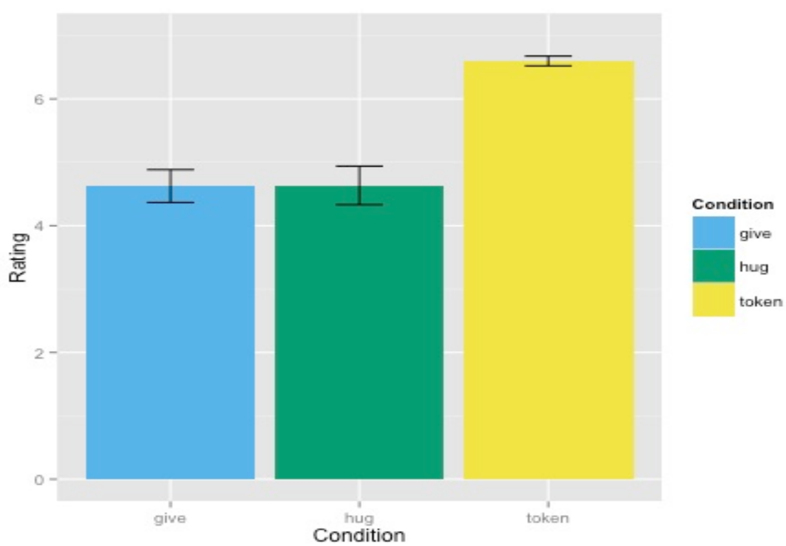
4.4.6 Predictions

Since *token comparisons* compare video tokens of the same type, the highest rating score (most similar) for this condition is expected. Apart from this, if giving-type contrasts are viewed differently from hugging-type contrasts, we expect a difference either in rating scores or in RTs, or both; if, however, the two contrasts are viewed similarly (for example, both contrasts merely involve a difference in the teddy-bear's presence/absence), then we expect no significant difference between the two conditions in either measure. If the two contrasts are viewed differently, specifically, for rating scores, we expect a higher score for hugging-type comparisons than giving-type ones, because the teddy bear in huggings is likely not as important as it in givings; for RTs, however, we do not have a prediction on the particular direction of difference, because a psychological important difference may result in longer time to respond to (effect of surprise), or it may as well take shorter (effect of salience).

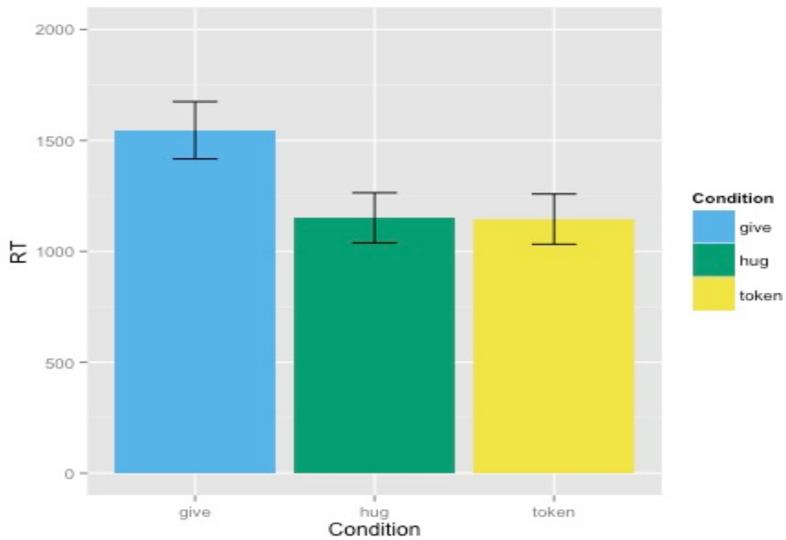
4.4.7 Results

Rating scores and RTs were analyzed separately. For rating scores, it is apparent that the token comparisons, not surprisingly, elicited the highest overall rating score. We conducted planned comparison between the *giving-type* and *hugging-*

type, using a one-tailed paired t-test, and found the mean rating score of giving-type (M=4.63, SD=0.90) is not significantly lower than that of hugging-type (M=4,64, SD=1.05), t(11) = -.011, p = 0.46). For RTs, we conducted planned comparison between the two types of contrasts, using a two-tailed t-test, and found that the mean RT of giving-type (M=1546.03ms, SD=447.39ms) is significantly longer than that of hugging-type (M=1151.09ms, SD=391.07ms), t(11) = 2.0276, p = 0.07. These results were plotted on Figure 4.13(a) and 4.13(b) respectively.



(a)



(b)

Figure 4.13: Mean ratings (a) **and RTs** (b) **by condition in Experiment 1**

4.4.8 Discussion

Lack of difference in the rating score suggested that the asymmetry we are interested in was not reflected in the explicit measure. This is not so surprising. There are many ways in which two things can be similar or different. And certainly, for both of our scene-types, the hugging and the giving, the gross difference between a video with and without a teddy bear is substantial, in contrast to the difference between tokens of the same type. And yet, although subjects assigned these contrasts similar rating scores, we did find an informative difference: judging by their *response times*, the similarity judgments were not made with equal ease between the two scene-types. It took subjects more time to make their similarity judgment for giving- than for hugging-type contrasts. This longer time taken in giving-type comparisons may be attributable to the *surprise* involved in not seeing an element that, because it fills a privileged role, is normally expected for scenes that fall under the concept GIVE. For the hugging-type comparison, however, if the teddy bear does not fill a role required for the concept HUG, then its absence should elicit no surprise.

4.5 Experiment 2-4: Adults, Number-Mismatch

We have discussed that concepts such as GIVE and HUG are typically expressed using sentences in which, plausibly, the number of syntactic arguments matches the number of participant roles. We now turn to the cases in which there is a potential mismatch between syntactic and conceptual information.

In Experiments 2-4, we extended our method to cases that we think people will represent under a three-participant concept. We want stimulus scenes that can be described using sentences like those in (4-4) – (4-6), repeated below, and are likely to be viewed, whether by infants or by adults, in ways that represent the roles of the lever (4-4), the victim (4-5), and a projectile (4-6) as participant roles, despite their not being required syntactic arguments for the relevant verbs.

(4-4) Anne *stole* a toy.

(4-5) Anne *jimmied* the box.

(4-6) Anne *beaned* Betty.

4.5.1 Participants

In each of Experiment 2-4, we recruited twelve adult participants from the University of Maryland undergraduate community, who were either given course credit or monetary reward for their participation.

4.5.2 Design

Experiment 2-4 used the same method as Experiment 1. In each experiment, there were two important comparisons. *Critical comparisons* contrasted two scenes where there is, potentially and plausibly, a difference in participant structure, like the contrast between Figure 4.10(a) and 4.10(b): in one, an object is plausibly involved as a participant in an event (e.g., a lever is involved as an instrument in opening a box); in the other, the same object is present but inert. *Perceptual comparisons* were between two scenes with no such difference, like the contrast between Figure 4.10(a) and 4.10(c): the second video merely reverses the orientation of the action in the first. Along with these two comparisons, there were also trials of *Token comparisons*, which were between two tokens of the same type of event. These three conditions were within-subject. Specific stimuli used in each experiment are introduced in Section 4.5.4.

4.5.3 Measurements

As in Experiment 1, we took participants' rating scores as the explicit measure, and the time taken to make a rating decision as an implicit measure. Rating scores were measured for all trials, but reaction times (RTs) were only measured on trials without replay.

4.5.4 Stimuli

In Experiment 2, one of the videos in a pair features a girl opening the box with a big hook. In the *critical* comparison for this experiment, the orientation of the action is the same, but the girl is using her hand to open the box, while the hook is held passively in her other hand. In *perceptual* comparisons, she opens the box with the hook, but she does so from the opposite side. See Figure 4.10 in Section 4.3.

In Experiment 3, one video shows a girl picking up a toy truck from out of another girl's hands. In *critical* comparisons, the other video shows the second girl simply standing by, while the toy is picked up from the table in front of her. Here, she is just a bystander. In *perceptual* comparisons, the second girl participates in both videos, just the orientation of the action is reversed. See Figure 4.14.

In Experiment 4, the base case features one girl hitting another girl by throwing a blue ball. In *critical* comparisons, the first girl again hits the other, but this time the ball does not leave her hand. The ball is not acting as a projectile. In the *perceptual* comparisons, again just the orientation of the action is reversed. See Figure 4.15.

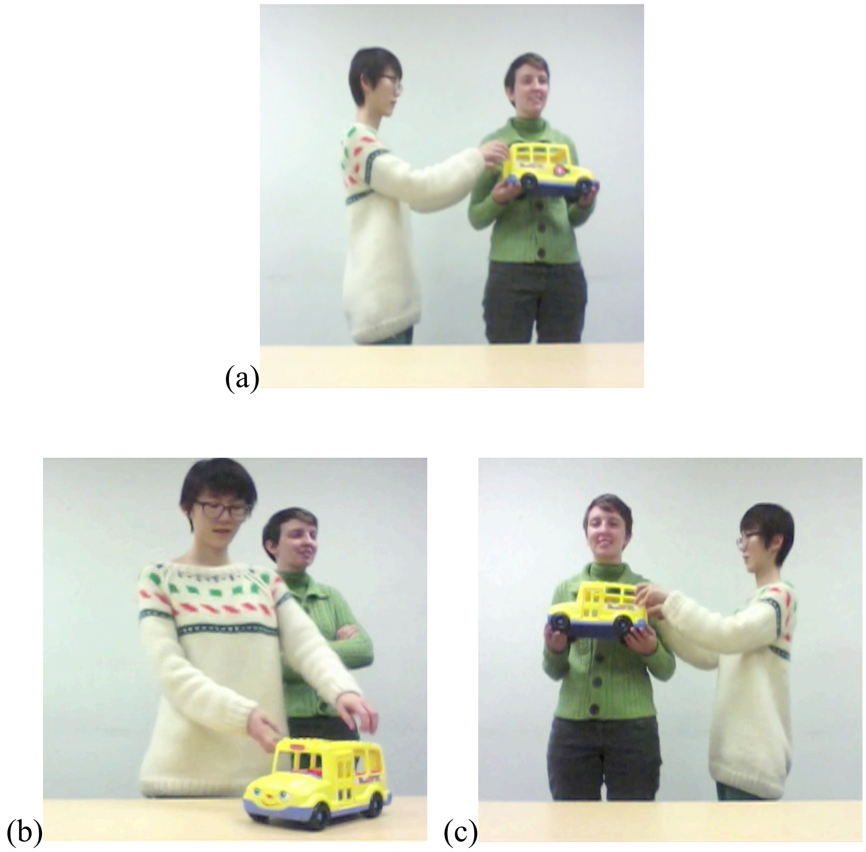


Figure 4.14: Stimuli in Experiment 3

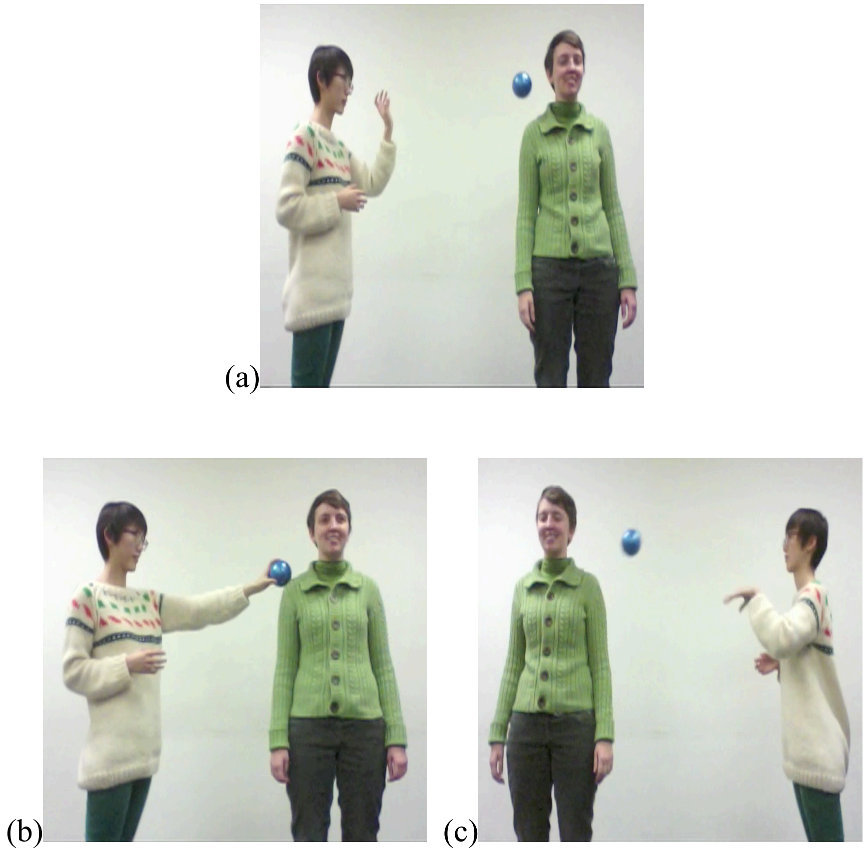


Figure 4.15: Stimuli in Experiment 4

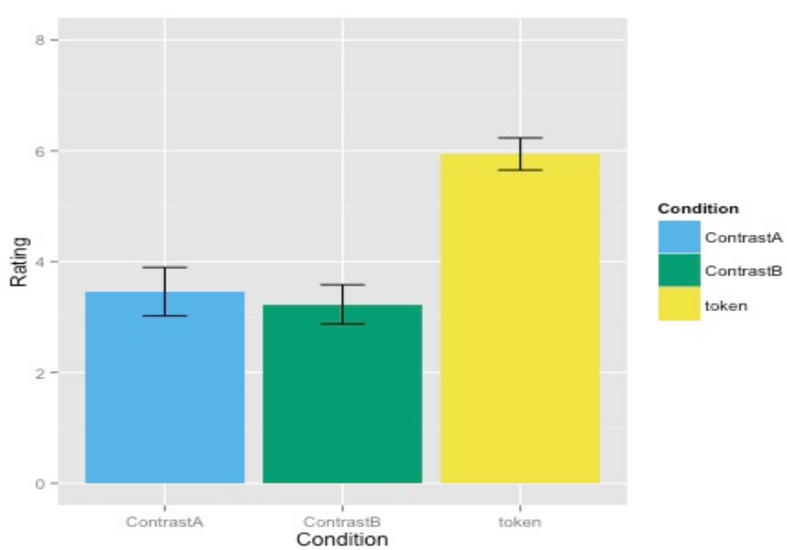
4.5.5 Predictions

If the tool in Experiment 2, the second girl as victim in Experiment 3, and the ball's role as a projectile in Experiment 4, are represented as important to the first scene in each experiment (scene (a) in Figure 4.10, 4.15 and 4.16), then adults should find the critical comparison less similar than they do the perceptual comparison, reflected in a lower mean rating score for critical comparisons, and/or a difference in mean RT between critical and perceptual comparisons (direction of difference unspecified).

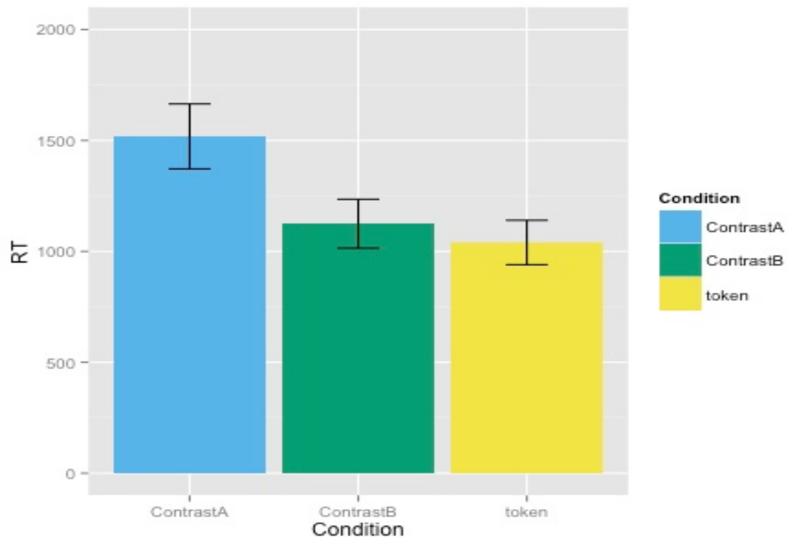
4.5.6 Results

For each of the experiments, it is apparent that the token comparisons elicited the highest overall rating scores. As for comparisons between critical and perceptual contrasts, I discuss results for Experiment 2-4 below.

In Experiment 2, a planned comparison between the rating scores in *critical* and *perceptual* trials, using a one-tailed paired t-test, revealed that the mean rating score of critical trials (M=3.46, SD=1.52) is not significantly lower than that of perceptual trials (M=3.23, SD=1.22), $t(11) = 0.96$, $p = 0.18$. For RTs, a planned comparison using a two-tailed t-test revealed that the mean RT of critical trials (M=1518.33ms, SD=507.13ms) is significantly longer than that of perceptual trials (M=1124.76ms, SD=381.67ms), $t(11) = 2.09$, $p = 0.06$. These results were plotted on Figure 4.16(a) and 4.16(b) respectively.



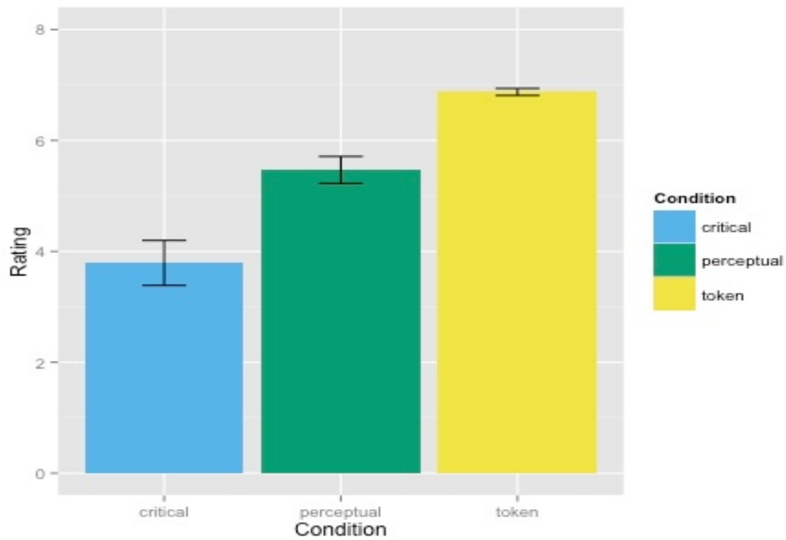
(a)



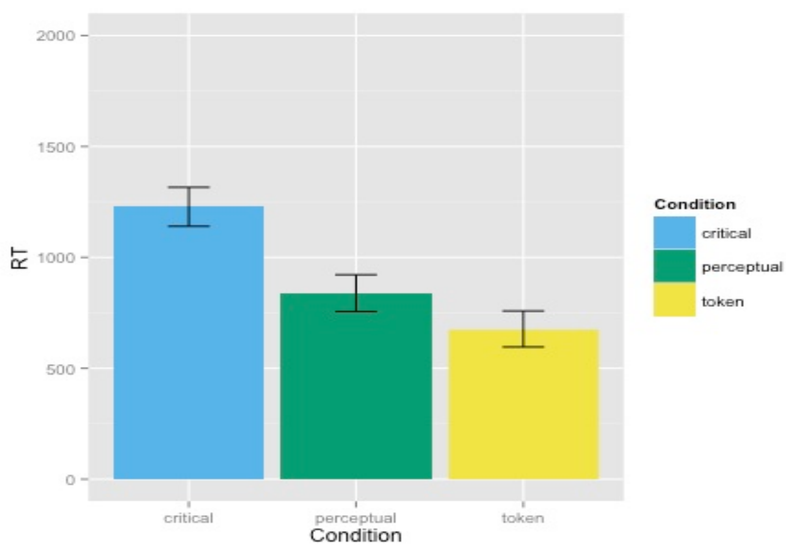
(b)

Figure 4.16: Mean ratings (a) **and RTs** (b) **by condition in Experiment 2**

In Experiment 3, like Experiment 1 and 2, we still found significant different in RTs between critical (M=1228.20ms, SD=304.49ms) and perceptual (M=838.94ms, SD=287.15ms) comparisons: $t(11) = 2.3987$, $p = 0.04$. But unlike Experiment 1 and 2, we also found significant difference in rating scores - $t(11) = -3.2612$, $p < 0.01$: perceptual comparisons (M=5.47, SD=0.84) were judged more similar than critical comparisons (M=4.07, SD=1.41). See Figure 4.17(a) and (b).



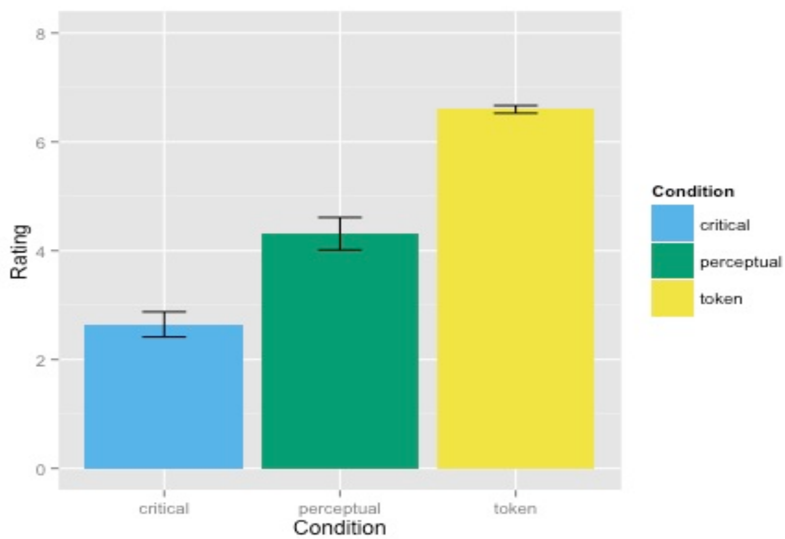
(a)



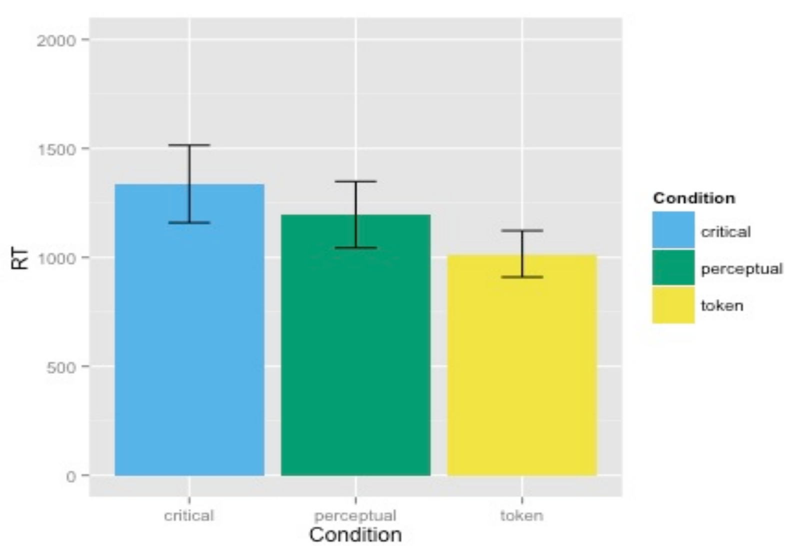
(b)

Figure 4.17: Mean ratings (a) **and RTs** (b) **by condition in Experiment 3**

In Experiment 4, unlike results found in the other experiments, however, here we found a difference between critical and perceptual comparisons only in rating scores - $t(11) = -5.1916$, $p < 0.01$: perceptual comparisons (M=4.31, SD=1.03) were judged more similar than critical ones (M=2.65, SD=0.80); no significant difference in RTs was seen - $t(11) = 0.7237$, $p = 0.48$: judgments for critical (M=1336.96ms, SD=615.43ms) and perceptual (M=1196.31ms, SD=528.34ms) comparisons were given at a roughly equal speed. See Figure 4.18(a) and 4.18(b).



(a)



(b)

Figure 4.18: Mean ratings (a) **and RTs** (b) **by condition in Experiment 4**

4.5.7 Discussion

A similarity judgment task with adults displayed sensitivity to the participation of a third entity in a variety of scene types. Sensitivity was revealed in either of two ways. For some type of stimuli (JIMMY/OPEN), the relevant measure was RT, as we found with GIVE/HUG in Experiment 1; for some, the relevant measure was rating score (BEAN/HIT); yet for some, both the explicit and implicit measures were revealing (STEAL/PICK-UP). We will not speculate about what accounts for the different expressions of sensitivity across scene types. For now, what is important is that we can feel confident in using these materials to investigate infants' nonlinguistic representation of participant-hood. With a clearer sense of this, we can ask what role participant representations play in verb learning, specifically with respect to PAM.

4.6 Experiment 5: Infants, Number-Match

With the results from the adult experiments (Experiment 1-4), we are confident that the materials we developed are appropriate for the purpose of assessing event perception. We therefore used them to probe event representation in prelinguistic infants. In this section I present Experiment 5, in which we successfully replicated Gordon (2003)'s findings about infants' perception of giving and hugging type scenes, using largely the same method with a few revisions.

4.6.1 Participants

Forty-eight English-speaking infants (24 boys, 24 girls) with a mean age of 10;00 (range: 09;17–12;13) participated in this experiment. Twenty-four additional infants were tested but excluded from the final sample because of experimental/equipment error (6), being unable to finish the experiment (10), parental interference (2), and failure to habituate (6). All infants were recruited through the Infant studies Consortium Date based at University of Maryland College Park.

4.6.2 Stimuli

The visual stimuli were the same pre-recorded events in Experiment 1. See the four types of events in Figure 4.11(a)-(d): a) giving-with-teddy, b) giving-without-teddy, c) hugging-with-teddy, and d) hugging-without-teddy. Each event lasted about 6 seconds and was repeated up to five times per trial, giving a maximum trial length of approximately 30 seconds. These stimuli were presented in complete silence, with no concurrent video.

If an infant looked away from the screen for more than 1 second, a video of a butterfly perched on a leaf (i.e. the attention-getter) was played until the infant's attention was recaptured. Another event – a flower bouncing event - was used at the end of the experiment as post-test to control for fatigue The duration of this event was about 15 seconds, and was played up to two times (i.e. 30 second maximum).

4.6.3 Apparatus

The experiment was run in the Habit version 1.0 program. The stimuli were played on a Samsung wall-mounted 51-inch plasma television, with built-in speakers, located 66 inches away from the chair (or highchair) where the infants were seated. A Sony EVI-D100 video camera was placed directly above the TV monitor. The camera was connected to a color TFT LCD monitor to allow the experimenter to observe the infant's eye fixation to the screen from a different room, and conduct online coding. Additionally, the video of the child, with a picture-in-picture display of what was on the TV screen, was captured on an iMac computer using QuickTime.

4.6.4 Design

Following Gordon (2003), this experiment used the Habituation-Switch Paradigm (Casasola & Cohen, 2000; Fennell & Werker, 2003; Werker et al., 1998; Younger & Cohen, 1986).

In the **habituation phase**, infants were presented trials of a particular event type for a maximum of 12 times (i.e. a maximum of 12 trials in this phase); they went through the trials until they reached a pre-set *criterion of habituation*, whichever came first. In this experiment, the criterion of habituation, was when an infant's average looking time during any block dropped to less than 50% of average looking time of the most-attended block (i.e. the block that has the longest total looking time); any three consecutive trials made a block. Therefore, the total number of habituation trials each infant received was different. These trials were randomized by blocks of three. Infants who did not meet the criterion of habituation were excluded from the sample of analysis, classified as exclusion due to *failure to habituate*. When infants reached the criterion of habituation or when the 12 trials were all played, whichever came first, the habituation phase was stopped and the **test phase** began. At test, all infants were presented a fixed number of 2 trials. These two trials were either the same event type as in habituation (*Same condition*), or a different event type (*Switch condition*). Following these two test trials was the **post-test phase**, where one post-test trial – the flower bouncing event - was presented. The purpose of having the post-test trial was to control for fatigue: if infants were still involved in the experiment towards the end (habituation but not fatigue), we would expect their attention to recover upon seeing the post-test, which was perceptually very distinct from the habituation and test trials.

In this experiment, infants were randomly assigned into two groups: in one group, they watched the giving-with-teddy type of event during habituation; in the other, they watched the hugging-with-teddy type. In each group, infants were further randomly assigned two conditions, which differed in what event type was played at test: in the Same condition, a different token of the type of event was played; in the Switch condition, the without-teddy event was played. See Figure 4.19 for an illustration of the design.

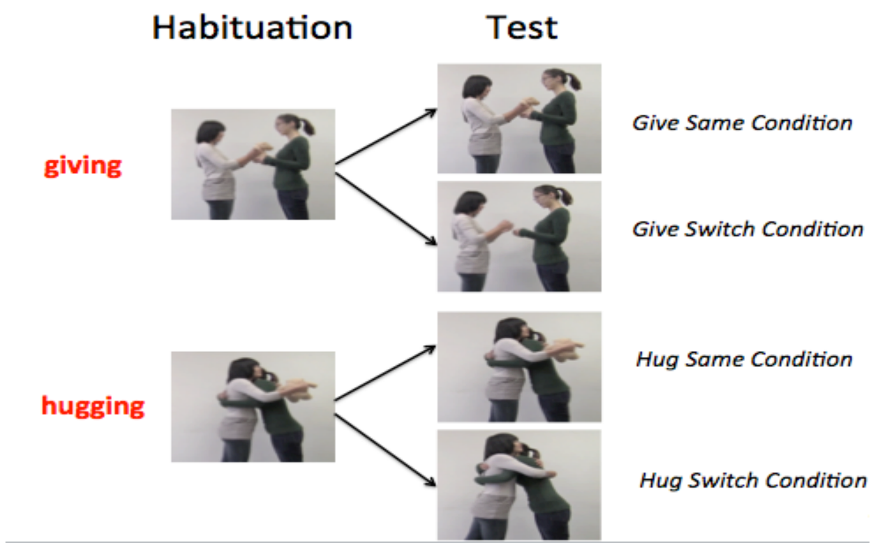


Figure 4.19: illustration of design of Experiment 5

Our design was different from Gordon (2003)'s in three aspects. First, Gordon used looking time *per trial* to decide habituation, but we used average looking time *per block* following the tradition of research using this paradigm, to avoid false habituation (i.e. low attention not because of habituation but due to accidental distraction). Second, Gordon did not use a post-test to control for fatigue; we did. Third, condition (same vs. switch) was used as a within-subject dependent variable in Gordon's study, but we used it as a between-subject variable, for this reason: in a within-subject design, when the Switch trial precedes the Same trial, event though Switch still has a switch, Same is no longer strictly speaking same, but could also be considered a switch, relative to the preceding Switch trial.

4.6.5 Procedure and Coding

The procedure began with obtaining the parent(s)' informed consent. When the infant was ready, he/she was led to the test room where the TV monitor and the digital camera were located. The parent came to the test room with the infant and stayed with him/her during the entire process. The infant sat either in the parent's lap or in a highchair in front of the monitor. We took precautions to ensure that the parent could not influence the child's behavior, by explicitly instructing the parent not to direct the infant's attention in any way, and asking the parent to wear a visor (to block sight) in cases where she chose to hold the infants on her lap.

The experimenter began the experiment in the control room next door, by setting up the computer to display an attention-getter (the butterfly). Once the infant looked at the attention-getter, the experimenter pressed the space bar on the computer to begin the first habituation trial. For each trial, the experimenter pressed a key on the computer when the infant attended to the screen, held the key for as long as the attention was maintained, and released it as soon as the infant looked away. A minimum of 2-second attention was required for it to be counted as a look. A trial continued until the infant looked away for more than 2 continuous seconds or until the end of the trial (approximately 30 seconds). The attention-getter came back on the screen to recapture the infant's attention at the end of each trial and stayed on the screen until the start of next trial. The experimenter was unaware of what phase of the experiment the child was in.

4.6.6 Measurement

The dependent variable for analysis was *looking time*, i.e. the amount of time spent on looking at the visual stimuli during a selected window. We used looking time to test for two things: a) successful *habituation* (controlled for fatigue), and b) *dishabituation* at test. To test for habituation, we compared average looking time of the first and last habituation blocks to see if there was a decrease in attention over the habituation trials. To control for fatigue, if infants' attention recovered upon seeing the post-test trial - indexed by a significant increase in average looking time from the last habituation block to post-test, we took that to mean habituation without fatigue. To test for dishabituation, we compared average looking time of the last two habituation trials with that of the two test trials; and we took significant increase of looking time from habituation to test as indicator of dishabituation.

4.6.7 Predictions

Infants in the Same condition were predicted to show no dishabituation at test, no matter they were in the giving group or the hugging group. Predictions for the Switch condition, however, vary depending on which type of event (givings or huggings) infants received: if in infants' view, the teddy bear in the giving event was 'psychological privileged' in some sense, but the one in the hugging event was not, then we would expect to see infants in the giving group dishabituate to the removal of the teddy bear but those in the hugging group do not.

4.6.8 Results

4.6.8.1 Habituation Controlled for Fatigue

To determine whether infants were successfully habituated, we conducted planned comparison between the first and last habituation block using one-tailed t-test, and found the mean looking time of the last block (M = 8.18s, SD = 4.02s) was significantly less than that of the first block (M = 18.87s, SD = 8.91s), $t(47) = 10.35$, $p < 0.001$. Thus, there was a significant drop in attention throughout the habituation phase. To make sure this habituation was not due to fatigue, we compared post-test to the last habituation block, as planned (one-tailed), and found that the mean looking time of the post-test (M = 19.08s, SD = 9.73s) was significantly greater than that of the last block (M = 8.18s, SD = 4.02s), $t(47) = 7.22$, $p < 0.001$. This showed that infants' attention recovered from habituation upon seeing the perceptually distinct post-test trial, suggesting they were not fatigued. These results were illustrated in Figure 4.20. Clearly, there was a clear decreasing in attention from the beginning (Hf1, Hf2 and Hf3 stand for first 3 habituation trials, respectively) to the end (Hl1, Hl2 and Hl3 stand for last 3 habituation trials, respectively, with Hl3 being the very last one); and attention recovered to some extent in post-test.

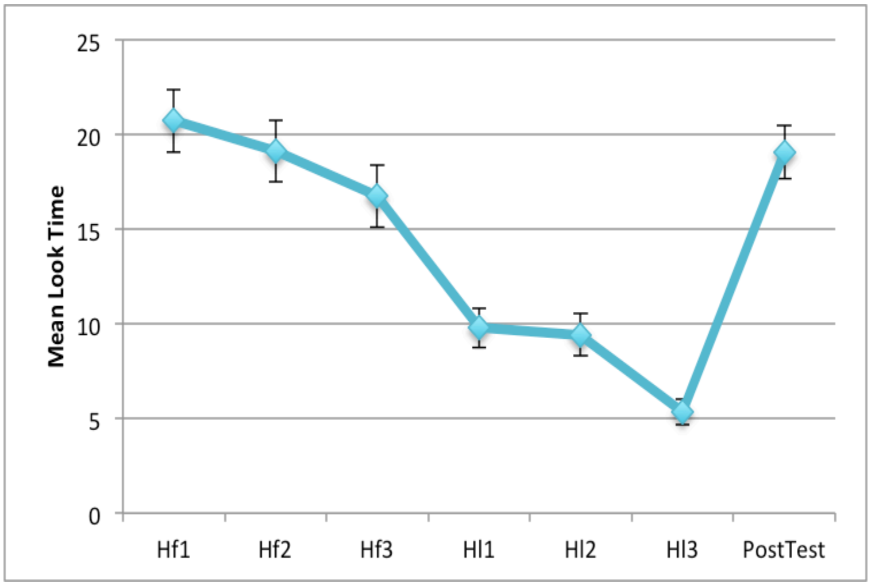


Figure 4.20: Habituation timeline of Experiment 5

4.6.8.2 Dishabituation Analysis

Having determined that infants were successfully habituated, we then conducted the main set of analyses on their performance at test. Data were first entered into a three-way mixed ANOVA with *event type* (givings vs. huggings) and *condition* (Same vs. Switch) as between-participants factors, and with *trial block* (last two habituation trials vs. two test trials) as a within-participants factor. This analysis revealed a significant three-way interaction, $F(1, 44) = 4.63$, $p = 0.04$.

To better interpret this interaction, data for each event type were analyzed separately, each with a two-way mixed ANOVA with *trial block* (last two habituation trials vs. two test trials) as a within-subject factor and *condition* (Same vs. Switch) as a between-subject factor. For the giving type of event, there was a main effect of condition, $F(1, 22) = 18.90$, $p < 0.001$; a main effect of trial block,

$F(1, 22) = 18.47$, $p < 0.001$; and a significant interaction between condition and trial block, $F(1, 22) = 10.12$, $p < 0.01$. Specifically, for the Same condition, attention during test (M = 7.01s, SD = 3.37s) was not significantly different from the last two habituation trials (M = 5.48s, SD = 3.11s): $t(11) = -1.56$, $p = 0.15$, two-tailed; but for the Switch condition, attention during test (M = 19.25s, SD = 8.83s) was significantly greater than that of the last two habituation trials (M = 8.96s, SD = 5.00s), $t(11) = -4.01$, $p = 0.001$, one-tailed. For the hugging type of event, there was no main effect of condition, $F(1, 22) = 1.75$, $p = 0.20$; no main effect of trial block, $F(1, 22) = 1.71$, $p = 0.20$; and there was no interaction between condition and trial block, $F(1, 22) = 0.09$, $p = 0.77$. Specifically, for the Same condition, attention during test (M = 7.9s, SD = 6.39s) was not significantly different from that during the last two habituation trials (M = 6.62, SD = 5.37s): $t(11) = -0.95$, $p = 0.36$, two-tailed; and for the Switch condition, attention during test (M = 10.46s, SD = 4.64s) was not significantly greater than that during the last two habituation trials (M = 8.45s, SD = 3.61s), $t(11) = -0.95$, $p = 0.18$, one-tailed.

See Figure 4.21 for illustrations of the results: trial blocks are plotted in blue and yellow bars, event types and conditions are plotted along the x-axis. Clearly, infants dishabituated only in the Switch condition of the giving type of event; the removal of teddy bear in the Switch condition of the hugging type of event did not cause the infants to dishabituate.

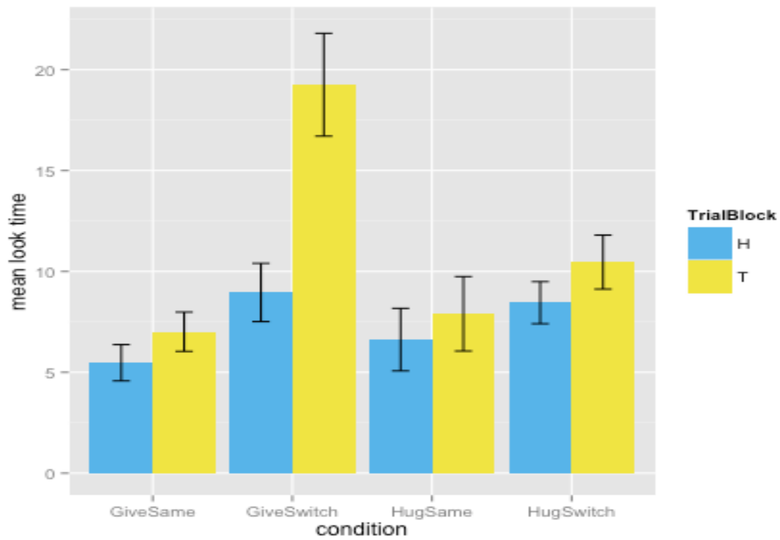


Figure 4.21: Mean look time across trial blocks in different conditions for different event types of Experiment 5

4.6.9 Discussion

Experiment 5, despite some small differences in design, successfully replicated Gordon (2003)'s finding that infants dishabituated to the disappearance of the teddy bear from giving type scenes, but not hugging type scenes. This suggests that infants do not view the hugging under a concept that has the role of the teddy as a participant role. The reason is, changes in a role that is psychologically foregrounded would presumably be noticed. Conclusions about the giving, where children do notice the disappearance of the teddy, are more delicate. There are many possible explanations for their noticing. The most plausible, we think, is that they view the scene under a concept that foregrounds the role of the object given. The two videos (hugging and giving) are objectively very similar; and yet they noticed the disappearance only in the giving video. A difference in the event representation is the best explanation of this difference.

4.7 Experiment 6-7: Infants, Number-Mismatch

With the method established by this number-match case (i.e. number of participants plausibly match number of arguments), we then moved onto the number-mismatch cases: jimmyings, stealings and beanings. By the time of this dissertation, we only obtained enough data on the jimmying case. In this section, I present two experiments that showed infants, like adults, viewed the lever used as a tool in box opening as important, independent of whether the lever's participant-hood is substracted (i.e. from active to inert) – Experiment 6, or added (i.e. from inert to active) – Experiment 7.

4.7.1 Participants

Twenty English-speaking infants (11 boys, 9 girls) with a mean age of 11;06 (range: 09;25 – 12;15) participated in Experiment 6; twenty-four (12 boys, 12 girls) with a mean age of 11;05 (range: 09;27 – 12;11) participated in Experiment 7. Eleven additional infants were tested but excluded from the final sample due to parental interference (1), being unable to finish the experiment (6), and and failure to habituate (4). All infants were recruited through the Infant studies Consortium Date based at University of Maryland College Park.

4.7.2 Stimuli

The stimuli in Experiment 6 were the same pre-recorded events in Experiment 2. See the three types of events in Figure 4.10(a)-(c): a) opening from left, lever active (i.e. jimmying), b) opening from left, lever inert, and c) opening from right, lever active. The stimuli in Experiment 7 were slightly different: a) opening from left, lever inert, b) opening from left, lever active, and c) opening from right, lever inert. Each event lasted about 6 seconds and was repeated up to five times per trial, giving a maximum trial length of approximately 30 seconds. These stimuli were presented in complete silence, with no concurrent video. Stimuli for the attention-getter and the post-test trial were the same as Experiment 5.

4.7.3 Apparatus

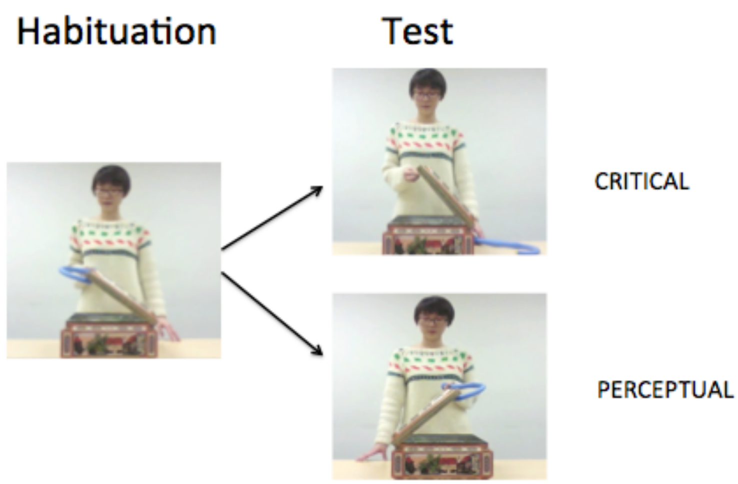
The same apparatus was used as in Experiment 5.

4.7.4 Design

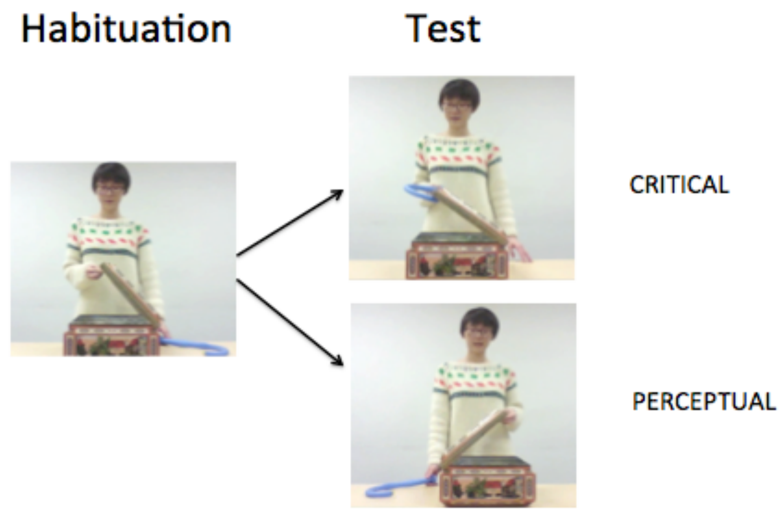
The overall structure was the same as in Experiment 5.

In Experiment 6, infants were presented a jimmying event until they reached the habituation criterion; then, at test, half of the infants who were randomly assigned to the *Perceptual condition* were shown a jimmying event from the opposite orientation, whereas the other half who were randomly assigned to the *Critical condition* were shown an opening event (same orientation as in habituation, lever inert); all infants then received one post-test trial. See Figure 4.22(a) for an illustration.

In Experiment 7, infants were habituated to an opening event; then at test, those in the Perceptual condition were shown an opening event form the opposite orientation, whereas those in the Critical condition were shown a jimmying event (same orientation as in habituation, lever active); all infants then received one post-test trial. See Figure 4.22(b) for an illustration.



(a)



(b)

Figure 4.22: Illustration of design of Experiment 6 (a) **& Experiment 7** (b)

4.7.5 Procedure and Coding

The procedure and coding are the same as in Experiment 5.

4.7.6 Measurement

As in Experiment 5, we used looking time as dependent variable to test for habituation and dishabituation.

4.7.7 Predictions

We predicted that infants in the Perceptual condition would show no dishabituation at test, but those in the Critical condition would dishabituate, if they viewed the lever actively involved in an opening as psychologically privileged in some sense.

4.7.8 Results

4.7.8.1 Habituation Controlled for Fatigue

Infants in Experiment 6 were successfully habituated, as revealed by a planned comparison between the first and last habituation block using one-tailed t-test: the mean looking time of the last block ($M = 4.92s$, $SD = 1.65s$) was significantly less than that of the first block ($M = 9.94s$, $SD = 4.16s$), $t(19) = 7.10$, $p < 0.001$. In addition, this habituation was not due to fatigue, because the mean looking time of the post-test ($M = 16.27s$, $SD = 4.62s$) was significantly greater than that of the last block ($M = 4.82$, $SD = 1.65s$), $t(19) = 10.83$, $p < 0.001$, one-tailed. See Figure 4.23 for habituation timeline.

Similarly, infants in Experiment 7 were also successfully habituated, as revealed by a planned comparison between the first and last habituation block using one-tailed t-test: the mean looking time of the last block ($M = 4.28s$, $SD = 1.77s$) was significantly less than that of the first block ($M = 9.77s$, $SD = 3.45s$), $t(23) = 10.97$, $p < 0.001$. And this habituation was not due to fatigue, because the mean looking time of the post-test ($M = 15.11s$, $SD = 4.10s$) was significantly greater than that of the last block ($M = 4.28s$, $SD = 1.77s$), $t(23) = 12.42$, $p < 0.001$, one-tailed. See Figure 4.24 for habituation timeline.

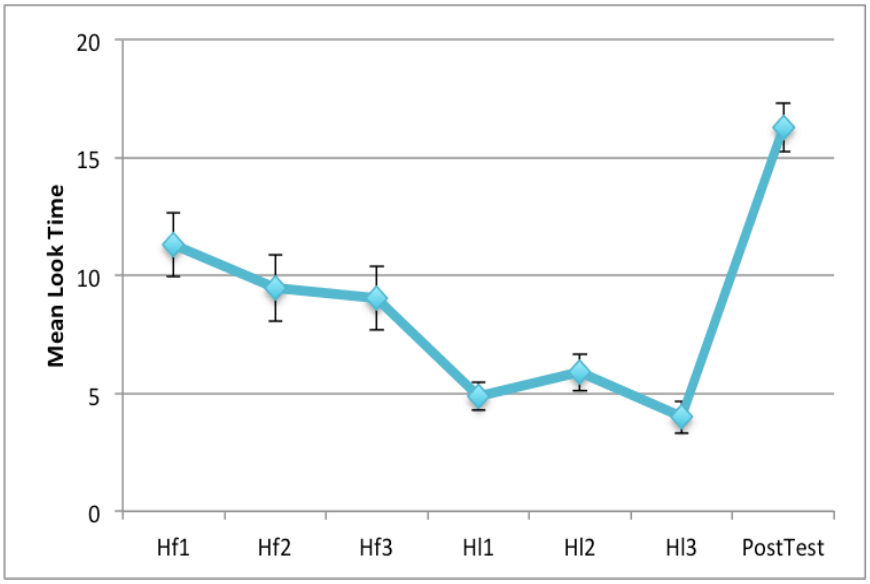


Figure 4.23: Habituation timeline for Experiment 6

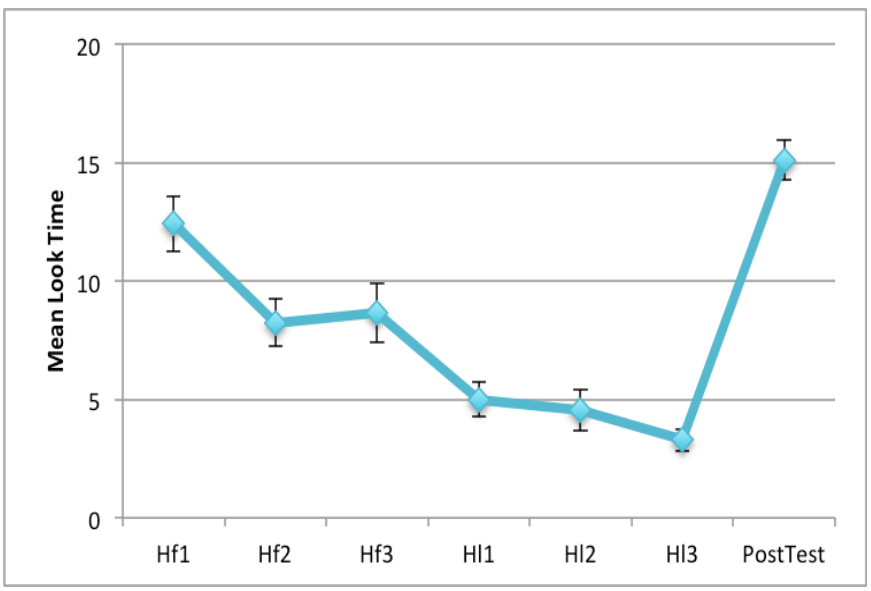


Figure 4.24: Habituation timeline for Experiment 7

4.7.8.2 Dishabituation Analysis

Data from Experiment 6 were entered into a two-way mixed ANOVA with *condition* (Perceptual vs. Critical) as between-participants factors, and with *trial block* (last two habituation trials vs. two test trials) as a within-participants factor. This analysis revealed a marginally significant effect of trial block, $F(1, 18) = 3.85$, $p = 0.07$; and a marginally significant interaction between condition and trial block, $F(1, 18) = 3.53$, $p = 0.08$. To interpret the interaction, we then conducted a comparison between trial block for each condition. Results were: for the Perceptual condition, mean look time during test (M = 5.24s, SD = 2.37s) was not significantly different from that during the last two habituation trials (M = 5.05s, SD = 2.59s), $t(10) = -0.25$, $p = 0.80$, two-tailed; but for the Critical condition, mean look time during test (M = 7.75s, SD = 3.82s) was significantly greater than habituation (M = 4.81s, SD = 1.91s), $t(8) = -2.20$, $p = 0.03$. These results suggested that infants did not dishabituate in the Perceptual condition, but did dishabituate in the Critical condition. See Figure 4.25 for illustrations of the results.

A similar analysis was conducted for Experiment 7, and revealed a main effect of trial block, $F(1, 22) = 12.56$, $p < 0.01$; and a marginally significant interaction between condition and trial block, $F(1, 22) = 3.61$, $p = 0.07$. To interpret the interaction, we then conducted a comparison between trial block for each condition. Results were: for the Perceptual condition, mean look time during test (M = 6.09s, SD = 4.36s) was not significantly different from that during the last two habituation trials (M = 4.45s, SD = 2.74s), $t(11) = -1.06$, $p = 0.31$, two-tailed; but for the Critical condition, mean look time during test (M = 8.85s, SD = 3.84s) was significantly greater than habituation (M = 3.40, SD = 1.88s), t(11) = -4.317, $p < 0.001$. These results suggested that as in Experiment 6, infants in this experiment also did not dishabituate in the Perceptual condition, but did dishabituate in the Critical condition. See Figure 4.26 for illustrations of the results.

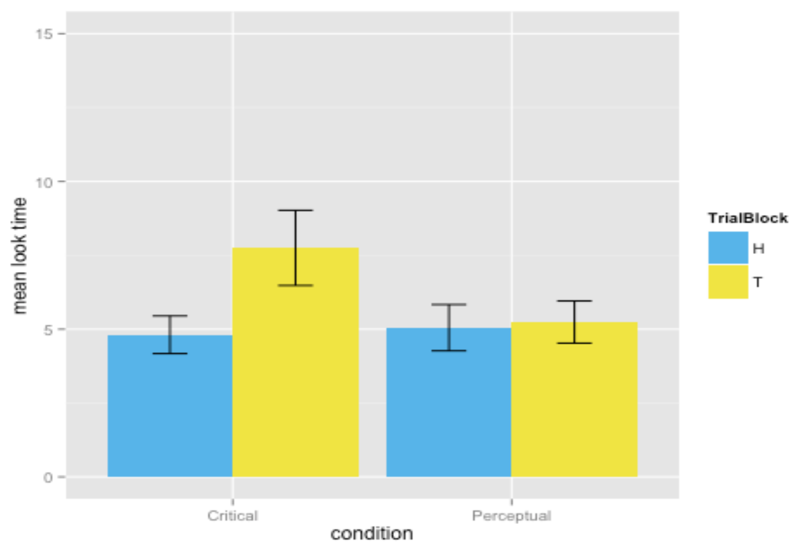


Figure 4.25: Mean look time across trial blocks in different conditions of Experiment 6

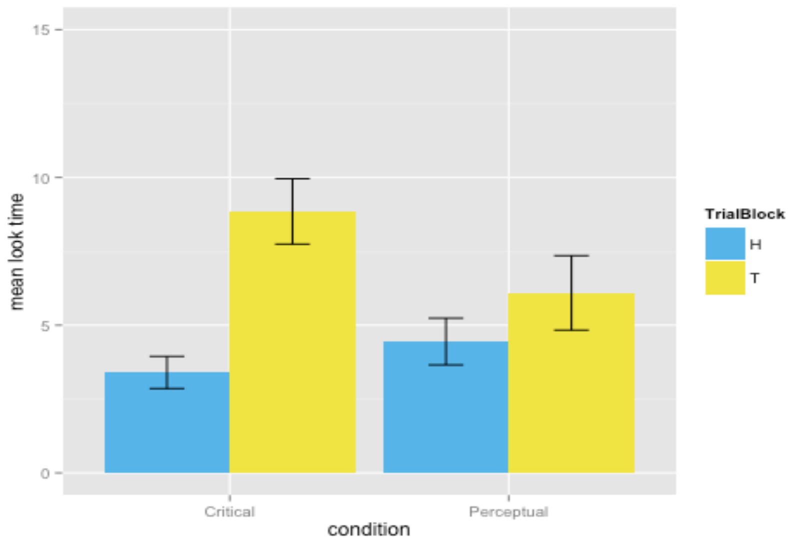


Figure 4.26: Mean look time across trial blocks in different conditions of Experiment 7

4.7.9 Discussion

These results suggest that infants, like adults, viewed these scenes under a concept that distinguishes the role of the lever as a participant; and they did so independent of whether the lever's participant-hood is subtracted or added. In their mental representation of these scenes, the lever is an explicit constituent, just like the eyes in our representation of a smiley human face. If this suggestion is correct, then the events of jimmying in our stimuli are viewed under a three-participant concept. And yet, sentences we normally use to describe such scenes will often have only two arguments. PAM therefore predicts that children should have difficulty learning verbs like *jimmy*.

4.8 General Discussion

To summarize, this chapter began with a review of previous literature that has lent evidence for the Participant-Argument-Match (PAM) hypothesis: it has been consistently reported that children's choice between a one-participant event (presumably so) and a two-participant event (presumably so) was modulated by the sentence structure they received; in particular, two-argument sentences directed their attention to the two-participant event, whereas one-argument sentences did not. These investigations, however, are limited in two ways, to the extent of being an adequate test for PAM: first, a complete test of PAM should go beyond 1- or 2-participant events and include events that are plausibly viewed under 3-participant concepts (e.g. givings, stealings); and second, for PAM to be testable, we must be confident about which concept the learner views a certain given scene under – in particular, in the learner's view of a certain given scene, how many participant roles are represented.

Motivated by these limitations, we conducted the current study to investigate the learner's event representations with a series of scene types that are plausibly viewed under three-participant concepts, including a case where the scene's descriptive sentence has the same number of arguments as its participants – givings; and a case where the scene is naturally described with a sentence that has no argument for a plausible participant – jimmyings, stealings, and beanings. As our preparatory studies to select the appropriate scene types to test with infants, we first tested adults, with some methodological attempts that generated interestingly revealing results. Our current results with infants, although only limited to givings and jimmyings, have important implications for the PAM hypothesis. Below I discuss these implications.

4.8.1 Implications on the Central Question

The goal of this chapter is to develop an adequate test for PAM, by fixing the extralinguistic side of the participant-argument-match equation.

The hypothesis for testing is: the learner uses the number of arguments to guide her interpretation of the verb. Quite straightforwardly, a test for this hypothesis would be: see if the learner's choice among multiple events that differ in the number of participants involved is modulated by the number of arguments in the concurrent sentence. In such a test, PAM predicts hearing an n-argument sentence would direct the learner's attention to an event she views as having event m participants ($m = n$), not to an event she views as having k participants ($k \neq n$). Such tests usually presume the ways under which the learner views the events presented to them as stimuli. The current study tested such presumptions.

Results from the current study suggest that infants are likely to view giving type of events and jimmying type of events as having 3 participants, but are likely to view hugging type of events and opening type (i.e. lever inert) of events as having 2 participants. With these independently obtained results, we could use these events as stimuli in our test for PAM. For example, to test if hearing a 2-argument sentence directs attention more towards a 2-participant event over a 3-participant one, we could design a preferential looking task as the following: we present on opposite sides of the screen a jimmying event as our 3-participant event and an opening-without-lever event as our 2-participant event, both having the lever present, and give half of the children a 2-argument sentence like 'Anne *gleebed* the box' and the other half a 3-argument sentence like 'Anne *gleebed* the box the lever'. PAM predicts that those hearing the 2-argument sentence would attend more to the opening event, whereas those hearing the 3-argument sentence would attend more to the jimmying event. The children might actually perform as predicted by PAM, which adds to the evidence supporting PAM.

But we soon find ourselves in a dilemma. On one hand, we are happy to see PAM's predictions are supported, even in cases of n = 3, adding to the existing evidence obtained from cases on n = 1 or 2. On the other hand, however, such results would suggest verbs like *jimmy* cannot be acquired under the guidance of PAM. Why so? The verb *jimmy* is not often used in 3-argument sentence as in 'Anne jimmied the box with the lever', because the lever is implied by the verb itself; rather, sentences with the verb *jimmy* is likely to occur in the input in 2-argument structure, like 'Anne jimmied the box'. A PAM-guided learner, hearing 'Anne jimmied the box', unfamiliar with the verb *jimmy*, will look for a 2-participant event as a candidate concept for the verb, but the real jimmying event in the world is viewed by the learner under a 3-particpant concept, as our experimental results suggest.

By the same logic, such problems may as well hold for verbs like *steal*, *rob*, *bean*, *sell*, *buy*, etc., each of which plausibly picks out an event that has more participant roles than its sentence has arguments. For example, *steal* and *rob* each picks out an event that necessarily has a thief, his loot, and a victim, otherwise not stealing or robbing, but simply picking-up; but sentences describing such events naturally occur in two-argument frames, like 'Anne stole the toy' or 'Anne robbed Betty'. Similarly, *bean* picks out an event that has a hitter, a hittee, and a thrown object, otherwise not beaning, but merely hitting; but its descriptive sentence is likely to be a two-argument one, like 'Anne beaned Betty'. Verbs like *sell* and *buy* were not discussed, but present a similar problem: *sell* or *buy* picks out an event that necessarily has a seller, a buyer, a property transferred, and transfer of money, otherwise not selling or buying, but simply giving or receiving; yet their descriptive sentences are likely to be 'Anne sold Betty a car', 'Anne sold a car', 'Betty bought a car from Anne', or 'Betty bought a car', all of which have fewer arguments than plausible participants.

In all these cases, PAM does not seem to be a reliable way to guide inference about the meaning of the verb, if independent evidence shows that the learner views these events in ways that have more participants than their descriptive sentences have arguments. Our investigation suggests that adults are likely to view stealing and beaning events with three participants. If these results hold in prelinguistic infants, as they do in the giving/hugging and jimmying/opening cases, then, we have reason to believe that these verbs are hard to acquire for a PAM-guided learner. Worse, PAM may mislead the learner to arrive at wrong hypotheses about these verbs – for example, taking *jimmy* to mean OPEN, *steal/rob* to mean PICK-UP.

Therefore, it seems the PAM hypothesis is challenged, especially in explaining how verbs like *steal*, *jimmy*, *bean*, are acquired. PAM may still hold as a valid learning guide, if some amendment/condition is added. Consider the following.

The first possible amendment is to impose some kind of restrictions on PAM to specify its scope. This could be done by either limiting its scope to the 1 and 2 cases, completely excluding cases of >=3. There are, however, two problems with this amendment. First, PAM limited to cases of 1 and 2 is indistinguishable from other heuristics like 'transitive/intransitive verbs name causative/non-causative events', which was one of the motivations for the current investigation on cases of 3. Second, it is not clear what 'excluding cases of >=3' means from the learner's perspective: If it means sentences with more than three arguments (n>=3) are beyond the scope of PAM, then the learner would not use PAM on sentences like 'Anne gave betty a teddy', but sentences like 'Anne stole a toy' would still remain in the set of sentences she would use PAM on, which does not solve the problem at all. If it means events with more than three participants (m>=3) are beyond the scope of PAM, then the learner would exclude events that she views as having 3-participant from her hypothesis space of possible verb meanings; from the data presented in Chapter 4, we know that learners plausibly view scenes of jimmying, stealing and beaning under 3-participant concepts, but these seem to be the target concepts of the to-be-acquired verbs; therefore, excluding these concepts only seems to worsen the problem. Hence, this possibility may not work after all.

Another way to maintain PAM as a learning guide but also account for the acquisition of verbs like *steal*, *jimmy* and *bean*, is to argue that these verbs are acquired after the consultation of PAM is over. Learning guides are only supposed to get learning off the ground and are expected to 'expire'. PAM, as an early learning bias, also has an expiration point. It is possible that these verbs are not learned, or mis-learned (e.g. misunderstood as more two-participant concepts like PICKUP, OPEN and HIT), during PAM's consultation; and when PAM no longer governs verb learning, these verbs can be learned, or previous misconceptions can be corrected. This idea necessarily admits that there might still be a period of 'misunderstanding', if PAM holds. Is there? Independent tests on learners' understandings of these verbs are required. Note that learners' production data may not be conclusive, because producing a word does not guarantee full comprehension of it. Additionally, we will need some independent account for how such misunderstandings are corrected – that is, how the learner later retreats from the wrong hypotheses they made under the guidance of PAM.

4.8.2 Some caveats

Our results from both adults and infants can only be taken to increase our confidence that the third entity is represented as an event participant, but are not conclusive. Participant roles are some subset of the roles entailed by an event predicate. For the notion to be used in language acquisition, this subset must comprise roles that are psychologically privileged in some way. The event concept must distinguish them psychological from the totality of what the concept entails. But until we know what factors constitute 'psychological privilege', we cannot be sure what participant-hood exactly is. Nevertheless, our attempts to experimentally approach people's representation of participants, with this working definition, is still of great value, because not every entailed element is equally represented, and psychological salience is clearly difference from perceptual salience.

Another caveat is that in experiments with adults, we took a significant difference between conditions in either the explicit measure (i.e. rating score) or the implicit measure (i.e. reaction time) to reflect adults' representation of participant-hood, but we did not have an explanation for why adults' representation was sometimes reflected in explicit judgment and sometimes in implicit reaction time. More experiments and replications with this method are needed. To the extent that either of these two measures revealed significant differences between conditions, this methodological attempt is still of importance.

4.8.3 Future directions

The current investigation points to a couple of directions for future work. First, the findings with adults on stealings and beanings set the stage for investigations with infants, which are underway in our lab. Second, a test for PAM on cases of n = 2 and 3 could be developed, now that we are more confident that givings and jimmyings are viewed under 3-participant concepts and huggings and openings under 2. Third, PAM predicts verbs like *jimmy*, *steal*, and *bean* may be hard to acquire; whether or not they are indeed hard to learn should be evaluated – this could be done by looking into child's production to identify the onset of these verbs in production, or by conducting a comprehension task of these verbs with children of different ages. In particular, PAM predicts these verbs may be misunderstood as some more general concepts that do not entail the third participant role, future work may need to a) show if there is indeed a stage of misunderstanding, and b) if so, provide an independent account for how such misunderstandings are corrected.

Chapter 5[27]

Cross-Linguistic Evidence on the Exact Nature of the Object-Names-Patient Expectation: Verb-Based versus Clause-Based

The Syntactic Bootstrapping Hypothesis (SBH) states that the child learner is able to utilize certain principled relations between verb syntax and verb semantics to make inference about verb meanings from syntactic cues (Gleitman, 1990; Landau & Gleitman, 1985). One aspect of such *principled relations* is the number match between syntactic arguments and event participants, characterized as the Participant-Argument-Match (PAM) bias, as discussed in Chapter 4. The current chapter discusses another aspect, which captures the correlation between the *syntactic position* of an argument and the *thematic role* the corresponding event participant takes; call it the position-role match aspect: for example, agents in the semantic structure are very likely to be subjects in the syntactic structure, and patients to be direct objects (Dowty, 1991; Fillmore, 1977; Grimshaw, 1981, 1990; Jackendoff, 1983; Jackendoff, 1987, 1990; Pinker, 1984a, 1984b, 1989a; Rappaport & Levin, 1988). If such correlations constitute part of the learner's early expectations, she could use such expectations to narrow her search space of

[27] The content in this Chapter was also reported in my PhD candidacy qualifying paper: Acquisition of Resultative Construction English- and Mandarin-learning children: a test case for participant-argument-matching. This paper was advised by Jeffrey Lidz, Alexander Williams, and Valentine Hacquard.

possible verb meanings down to those events in which the referent of the subject/object is the agent/patient. For example, given the sentence 'the bunny is *gorping* the duck', the learner guided by such an expectation will infer that the novel verb gorp names an event of which the bunny is an agent and the duck is a patient. In this chapter I focus on the correlation between the *direct object* position and the *patient* role, and examines a possible expectation the learner may hold: objects name patients, dubbed the *Objects Name Patients (ONP)* expectation.

In this chapter, I ask a new question about the exact nature of the ONP expectation: does the learner expect the object to be the patient of the event described by the *clause*, or, the event of the *verb*? This question is important because the event described by the clause may differ from that described by its main verb; for example, if the predicate is complex, comprising both the main verb V and some secondary predicate X, the events that satisfy V+X may differ from those that satisfy V alone. When they differ, it is crucial for the learner to know which event (thus, which patient) the ONP expectation refers to, so as to be able to use ONP correctly. This question, to my knowledge, has never been asked before. This is because most previous research examined simple-predicate clauses, a context where the distinction of interest does not matter: in a simple-predicate clause, there is only one verb, thus the verb's event and the clause's event are identical. Take the sentence 'Al pounded the cutlet' for example, the event described by the clause is a POUNDING event, and the event described by the verb *pound* is also a POUNDING event. In this chapter, I discuss a case where the distinction becomes particularly important – the case of complex-predicate clauses. Take the sentence 'Al pounded the cutlet flat' for example; here, the event described by the clause is a POUNDING-FLAT event, which is related, but not identical to, the event described by the verb, a POUNDING event.

The discussion will proceed in light of some observed cross-linguistic variation, which makes the distinction of interest – verb-based ONP versus clause-based ONP – event more intriguing. In some languages (e.g. English), the arguments correspond to the participants picked out by the verb's event – objects name patients of *the event of the verb*, whereas in some languages (e.g. Mandarin), the arguments correspond to the participants picked out by the clause's event – objects name patients of *the event of the clause*. In this chapter, I will present a preferential looking experiment that examines Mandarin-learning and English-learning toddlers' ONP expectations and asks whether they are sensitive to the language-specific properties with regard to the ONP expectation. Our data suggest that young leaners, respect ONP in general, but they may be entertaining different versions of ONP depending on language-specific properties; specifically, Mandarin-learners may entertain *a clause-based ONP* (i.e. expect the direct object to name the patient of the clause's event), whereas English-learners may entertain *a verb-based ONP* (i.e. expect the object to name the patient of the verb's event).

In what follows, I will begin with a review of previous literature supporting Syntactic Bootstrapping Hypothesis – the position-role match aspect of it in particular, and its cross-linguistic universality; and bring on the table the distinction between *event of the verb* and *event of the clause* in the context of complex-predicate sentences, as well as some cross-linguistic variations in English and Mandarin with respect to the Resultative Construction (Section 5.1). Then, I will discuss what implications the cross-linguistic variations may have for acquisition and for the learning bias under discussion – the ONP expectation, which motivate the current experiments (Section 5.2). Subsequently, I will present two experiments, with novel word learning tasks, looking at adults' and young toddlers' expectations about the thematic relations of the direct object to the verb in Resultative Constructions, comparing English and Mandarin population (Section 5.3 – 5.5). From there, I will discuss what the findings imply about the current acquisition theory and what could be done to perfect this theory (Section 5.6).

5.1 Background

This section first reviews supporting evidence for Syntactic Bootstrapping Hypothesis and its universality, pointing out two limitations of current literature – little research on complex-predicate sentences, and limited cross-linguistic examination on the position-role match aspect (Section 5.1.1), then discusses why the context of complex-predicate sentences is important (Section 5.1.2), and subsequently introduces a particular complex-predicate construction in light of cross-linguistic data that the current study will use as a test case (Section 5.1.3).

5.1.1 Literature: Evidence on Syntactic Bootstrapping

The Syntactic Bootstrapping Hypothesis has received important initial evidence. In Chapter 4, I reviewed some evidence that children learning English rely on the number-match aspect of this hypothesis to help them select from candidate events with different numbers of participants – specifically, a 2-argument sentence picks out a 2-participant event and a 1-argument sentence picks out a 1-participant event. Here I focus on two other aspects of the literature: first, children not only use the number of arguments but also use the syntactic position of arguments as cues to verb meaning; and second, evidence shows Syntactic Bootstrapping is a universal learning bias.

5.1.1.1 Position-Role Correspondences

The findings reviewed in Chapter 4, although also consistent with the hypothesis that children use syntactic positions to identify participant roles, did not specifically test it. For example, in Naigles (1990), children who heard the transitive sentence 'the duck is *gorping* the bunny' attended more to the event where the duck was pushing the bunny down; this was consistent with children mapping the subject NP to the agent and the object NP to the patient; but without a competing scene that reversed the thematic relations (e.g. an event where the bunny was pushing the duck down), it was still not conclusive.

The first conclusive evidence about young children's ability to use syntactic positions to infer thematic relations was from Hirsh-Pasek, Golinkoff, Fletcher, DeGaspe-Beaubien, and Cauley (1985). With a Preferential Looking Paradigm, the 17-month-old English-learning infants were presented two videos simultaneously - in one, Big Bird was tickling Cookie Monster; in the other, Cookie Monster was tickling Big Bird. Concurrent with the scenes, some infants heard the stimulus sentence 'Big Bird is tickling Cookie Monster', and some heard 'Cookie Monster is tickling Big Bird'. Results were: infants who heard the sentence with Big Bird as the subject and Cookie Monster as the object looked more to the event with Big Bird as the agent and Cookie Monster as the patient, and vice versa. These results suggested that infants this age clearly understood English word order (that subject precedes the verb and object follows the verb), and importantly, they understood the mapping between a particular syntactic argument and a thematic role of a participant (i.e. subject-agent, object-patient).

Fisher et al. (1994) investigated whether knowledge of position-role correspondences could help learners learn verbs. In particular, they looked at a set of verbs that were considered specifically problematic for observation-based learning - perspective-taking verbs, like *give* and *receive*, *chase* and *flee*, *ride* and *carry*, *eat* and *feed*, etc. Verbs in each pair are likely to pick out the same event, and which verb to use to label that event depends on which perspective is taken. To distinguish the verbs in each pair, information gathered from the number of argument NPs may not be sufficient, and information obtained from the positions of argument NPs may help. For example, hearing a sentence like 'the bunny is *zilking* the monkey' in concurrence with a scene where a bunny sits on the back of a monkey; parsing the sentence into a 2-argument structure is not going to help distinguish between ‘riding’ and ‘carrying’; understanding that the subject NP - ‘the bunny’ – names the agent of the event and object NP – ‘the monkey’ – names the patient, will help decide that the verb zilk means riding instead of carrying.

Fisher et al. (1994) tested this knowledge with 3- and 4-year-old English-learning children. Children were shown a scene (e.g. a bunny sits on the back of a monkey) and were presented a different sentence depending on the condition - e.g. ‘look, the bunny is *zilking* the monkey’, or ‘look, the monkey is *zilking* the bunny’, or ‘look, *zilking*’. They were then asked what the novel verb means (e.g. ‘what does *zilking* mean’). Results were: although children sometimes had a biased hypothesis about the novel word’s meanings when no argument was accompanying the verb, this bias was overcome when they heard a sentence providing opposite information. For example, they were biased to interpret *zilking* in ‘look, *zilking*’ as carrying; but if the sentence they heard was ‘look, the bunny is *zilking* the monkey’, then they more often answered ‘riding’ as the meaning of the novel verb. These results held across trials.

The above findings suggested that young learners were sensitive to the correspondence between syntactic positions of arguments and thematic roles of participants from a very young age (i.e. 17 months), and were able to use such knowledge to constrain their hypotheses about novel verbs at around 3 years.

5.1.1.2 Syntactic Bootstrapping as a Universal Learning Bias

This section reviews evidence supporting Syntactic Bootstrapping as a universal bias, most of which are on the number-match aspect of this hypothesis (i.e. PAM), because few has studies the position-role match aspect cross-linguistically.

In addition to the bulk of evidence from children learning English (Fisher, 1996; Fisher et al., 1994; Hirsh-Pasek et al., 1985; Hirsh-Pasek et al., 1996; Naigles, 1990; Naigles, Fowler, & Helm, 1992; Naigles & Kako, 1993; Yuan & Fisher, 2009; Yuan et al., 2012; inter alia), the Syntactic Bootstrapping Hypothesis has also received evidence from other languages (e.g. Naigles & Lehrer, 2002 on French). Importantly, evidence for this hypothesis also come from many argument-drop languages, for which a bias relying on counting argument number and identifying argument positions may be unreliable: e.g. Lidz et al. (2003) on Kannada; Göksun, Küntay, and Naigles (2008) on Turkish; Matsuo, Kita, Shinya, Wood, and Naigles (2012) on Japanese; and Lee and Naigles (2008) on Mandarin. Syntactic Bootstrapping, therefore, has been considered a universal learning bias of an unlearned origin.

Lidz et al. (2003) provided two arguments for Syntactic Bootstrapping being a universal unlearned bias. First, isolated deaf children who have no exposure to any language spontaneously invent home signs that demonstrate a similar mapping principle; specifically, invented gestures that include similar categories like nouns and verbs are combined into sentences that vary in the number and positioning of the noun phrases depending on the meanings of the verb (Feldman, Goldin--Meadow, & Gleitman, 1978; Goldin--Meadow & Mylander, 1984). Second, 3-year-old children learning Kannada, a language in which certain morphological marking reliably indicates causativity (with a morpheme similar to the English "-ize") whereas transitivity (the number of noun phrases) do not always have a causative meaning, still consistently rely on the less reliable predictor – number of noun phrases - to act out verb meanings; and the verbs are all known motion verbs for which there is a principled link between transitivity in structure and causativity in meaning.

Göksun et al. (2008) looked at children learning Turkish (2-, 3-, 4-, and 5-year-olds), a language that allows argument ellipsis and has inflectional morphology (e.g. accusative markers). They had participants act out transitive and intransitive verbs in four different frames: a) two-argument frame without accusative marker (NNV), b) two-argument frame with accusative marker (NNaccV), c) one-argument frame without accusative marker (NV), and d) one-argument frame with accusative marker (NaccV). For each enactment, causativity of the enacted event was coded. The authors found that children's enactments included more causative events when the sentences they received were in two-argument frames, and less when the sentences they received were in one-argument frames; and causative enactments also increased when they received sentences with accusative markers. These results suggested that Turkish-learning children, despite their input language drops arguments, still used the number of arguments as cues to verb meanings; and they also utilized morphological cues.

Lee and Naigles (2008) looked at Mandarin-learning children. The authors considered Mandarin the 'worst case scenario' for Syntactic Bootstrapping because it allows dropping of both the subjects and objects and it has no morphological cues to the verb argument structure. Despite this 'disadvantage', they demonstrated via an act-out task that 2- and 3-year-old children learning Mandarin still adhered to Syntactic Bootstrapping, entertaining a causative meaning for a familiar intransitive verb when a post-verbal NP is added, and entertaining a non-causative meaning for a familiar transitive verb when the post-verbal NP is deleted.

5.1.1.3 Summary

Taken together, Syntactic Bootstrapping seems to be a universal learning bias that guides children's verb learning via the principled relations between syntactic arguments and event participants. Children could use the number correspondence as well as the position-role correspondence to zoom their attention in onto some events but not others, and this seems to hold across a variety of languages.

The existing literature, however, is limited in two ways. First, previous research, no matter on the number-match aspect or the position-role match aspect, has been limited to simple transitive and intransitive sentences (i.e. simple-predicate sentences), leaving its generalizability to more complex structures untested, like complex-predicate sentences - one comprising two or more predicate words. The context of complex-predicate sentences is important for studying the exact nature of the Syntactic Bootstrapping Hypothesis, because a distinction that is not specified in the current theory – whether the participants are participants of the verb's event or participants of the clause's event – is only visible in this context. For example, the hypothesis states that direct objects name patients (ONP). Patient of which event is the hypothesis referring to – patient of the verb's event, or patient of the clause's event? Second, most cross-linguistic examinations of the Syntactic Bootstrapping hypothesis have focused on the number-match aspect of it, but few has examined children's expectations about the position-role correspondence. For example, the hypothesis states that children expect direct objects to name patients (ONP). Does ONP hold cross-linguistically? Taking the two aspects together, do children learning different languages expect the direct objects to name patients of the event; which event (verb's vs. clause's) do they expect, and does this vary cross-linguistically? I will discuss the distinction – verb's event vs. clause's event in Section 5.1.2, and some cross-linguistic observations in this regard in Section 5.1.3.

5.1.2 A Distinction – Verb-Based ONP vs. Clause-Based ONP

Usually in talking about the Syntactic Bootstrapping Hypothesis, hence the Object Names Patient expectation (ONP), it is implicitly assumed that the event described by the clause is exactly the event described by the verb. No distinction needs to be made, in all above reviewed experiments, because the predicate in the testing sentence is simple, comprising just a single verb. In simple-predicate clauses like ‘Al pounded the cutlet’, for example, the event of the verb and the event of the clause are identical, both the POUNDING event. Identical events certainly pick out identical sets of participants; from an ONP-guided learner’ perspective, therefore, the distinction does not matter.

However, not every sentence has a simple predicate. For complex-predicate sentences, take for example ‘Al pounded the cutlet flat’. This mentions at least two events – an activity of pounding and a state of being flat; it is possible that the clause also mentions a third event, a change brought about by the pounding and ending with flatness – a pounding-flat event. Here, the distinction emerges: the event of the verb is pounding, whereas the event of the clause is pounding-flat. Importantly, the pounding-flat event may not be identical to the pounding event, since one can be slow when the other is fast, as in the sentence ‘Al slowly pounded the cutlet flat, by pounding it rapidly for hours’ (Goldberg & Jackendoff, 2004; Rappaport & Levin, 2001; Williams, 2014). Despite their being different events, they may still share their participants; for example, the agent of pounding is the same person (i.e. Al) as the agent of pounding-flat, and the patient of pounding is the same entity (i.e. cutlet) as the patient of pounding-flat. In such cases, the distinction between the two types of events may still not matter from an ONP-guided learner’s perspective, because all that the learner cares about is the subject/object identifies the agent/patient of the event (pounding or pounding-flat). Nevertheless, sometimes the verb’s event and the clause’s event are not only distinct events, but also pick out different sets of participants. In such cases, as we will see in the case of Resultative Constructions in Mandarin (Section 5.1.3), the distinction becomes importantly different: from the learner's perspective, expecting the object to name the patient of the event is no longer sufficient, for there may be two events which have different patients. Then the question becomes: ONP states that direct objects name patients, but patient of which event – the verb's event or the clause's event?

In light of such cases, it may be necessary to specify which event the Syntactic Bootstrapping Hypothesis, hence the theory of ONP, is referring to – the verb's event or the clause's event. The relevant distinction is between a *verb-based SBH* (hence *verb-based ONP*), which is likely what most researchers in the literature are entertaining and thus the 'standard view', and a *clause-based SBH*[28] (hence *verb-based ONP*). See (5-1) and (5-2) for definitions. Verb-based SBH regulates the relation between verb meaning and verb valence; thus a verb-based ONP guided learner would expect the object in a sentence to name the patient of the main verb's event. On the other hand, clause-based SBH regulates the relation between sentence meaning and sentence valence; thus a learner guided by the clause-based ONP would expect the object in a sentence to name the patient of the clause's event.

(5-1a) **Verb-based SBH**: The clause has an argument for every participant of *each* event predicate within its verb phrase.

[28] Clause-based SBH was never really considered by lexcalists like Fisher, Gleitman etc., who assume a lexicalist architecture. But it may be explicitly considered by Kako and Wagner (2001) and by Goldberg (1995), though perhaps not precisely in these terms.

(5-1b) **Verb-based ONP**: The direct object of the sentence names the patient of the event of the main verb.

(5-2a) **Clause-based SBH**: The clause has an argument for every participant of the event predicate given by its verb phrase.

(5-2b) **Clause-based ONP**: The direct object of the sentence names the patient of the event of the clause.

As I mentioned, however, this distinction may still be unnecessary *if* the verb's event and the clause's event in complex-predicate sentences the learner receives always pick out the same set of participants. It is only when the two types of event pick out different sets of participants that this distinction becomes important. In the next section (Section 5.1.3), I will discuss such a case – Resultative Constructions in Mandarin.

5.1.3 A Cross-Linguistic Variation in Resultative Constructions

Resultative Constructions (RCs) consist of a main predicate denoting the main event, a secondary predicate denoting the result of the main event, and a direct object. Following Williams (2008), I will call the main predicate *the means predicate (M)* and the secondary predicate *the result predicate (R)*. Semantically, RCs express a causative relation between the events denoted by M and R, and the direct object in RCs is the entity that undergoes the result state defined by R. For example, in (5-3), the event denoted by M (i.e. wiping) causes the event denoted by R (i.e. cleanness), and the direct object (i.e. the table) enters the cleanness state as a result of M. In Williams (2008)'s terminology, the direct object NP *controls* R.

(5-3) She wiped the table clean.

English and Mandarin RCs have roughly the same structure, with a difference lying in the order of R and its controller DO (direct object) – in English, DO precedes R, whereas in Mandarin DO follows R[29]. The RC structures in English and Mandarin are schematized in (5-4a) and (5-4b) respectively (S: subject).

(5-4a) S – M – DO – R

(5-4b) S – M – R – DO

The crucial difference between English and Mandarin RCs that interests us is the argument requirements of the means verb[30]. In English, the patient of the

[29] In Mandarin, R immediately follows M, and M-R is sometimes analyzed as a complex predicate (Huang, 1992; Li, 1990).
[30] We will only focus on cases where the means predicate is a transitive verb; cases where it is an intransitive verb are not within our discussion (e.g. He cried his throat hoarse).

event denoted by the means verb must be syntactically realized, as the noun phrase in DO position. In contrast, in Mandarin, the patient of the means verb event does not have to show up; instead, it is the patient of the whole sentence (or, the resultative predicate M-R) that takes the DO position. Sometimes the patient of the means verb and the patient of the complex predicate happen to be the same entity in the world - for example, the thing being wiped and the thing being wiped-clean are both the table, while sometimes they correspond to distinct entities – for example, the thing being wiped is the table whereas the thing being wiped-dirty is the cloth. It is when the participant is *not* shared by the two events – *event of the verb* and *event of the clause* – that English and Mandarin show diverging grammatical patterns, as in (5-5b) and (5-6b). In particular, English requires the patient of the verb to always be present in the sentence, whereas Mandarin is more flexible, in the sense that as long as the patient of the sentence shows up, it does not have to be the patient of the means verb – it can be a non-patient role of the means verb's event, the instrument (e.g. the cloth) for instance.

(5-5a) She wiped the table clean. [(5-3), repeated here]

(5-5b) *She wiped the cloth dirty. [Intended meaning: she used the cloth as the tool of wiping something, and as a result, the cloth became dirty]

(5-6a) Ta ca ganjing le zhuozi.

She wipe clean ASP table

She wiped the table clean.

(5-6b) Ta ca zang le mabu.

She wipe dirty ASP cloth

She wiped something with the cloth, and the cloth turned dirty.

See (5-7) – (5-9) for more examples of Mandarin RCs: In all of the a sentences (5-7a, 5-8a, 5-9a), like (5-6a), the direct object NPs all refer to the patient of the main verb's event; whereas in all of the b sentences (5-7b, 5-8b, 5-9b), like (5-6b), all of them refer to the patient of the clause's event, but a non-patient role (i.e. in these cases the instrument) from the perspective of the verb's event.

(5-7a) Ta qie sui le rou.

He cut into-pieces ASP meat

He cut the meat into pieces.

(5-7b) Ta qie dun le dao.

He cut dull ASP knife.

He cut something with the knife, as a result the knife turned dull.

(5-8a) Ta ti duan le shuzhi.

He kick broken LE tree-branch

He kicked the tree branch and as a result the branch split.

(5-8b) Ta ti teng le jiao.

He kick hurt ASP foot

He kicked something with his foot, hurting his foot.

(5-9a) Ta hua hei le meimao.

She draw black ASP eyebrows

She drew her eyebrows black.

(5-9b) Ta hua tu le meibi.

She drew blunt ASP eyebrow-pencil

She drew something with the pencil, and as a result the pencil turned dull.

Importantly, even though in RCs, Mandarin allows the noun phrase in the direct object position to be interpreted as a non-patient role of the means verb's event, it does not allow such as interpretation in simple clauses, as in the contrast between (5-11a) and (5-11b). This is the same as in English; see the contrast between (5-10a) and (5-10b). This is not surprising, because the reason why Mandarin allows an instrument interpretation in RCs is because the instrument of the means event happens to be the patient of the complex predicate (i.e. the patient of the sentence); in simple clauses, however, the patient of the verb and the patient of the sentence are equivalent, thus disallowing any non-patient interpretation.

(5-10a) She wiped the table.

(5-10b) *She wiped the cloth. [intended meaning: wiping using a cloth]

(5-11a) Ta ca le zhuozi.

She wipe ASP table

She wiped the table.

(5-11b) *Ta ca le mabu.

She wipe ASP cloth

She wiped something with the cloth.

The patterns in RCs and simple transitive sentences in English and Mandarin are summarized in Table 5.1. In Williams (2008)'s terminology, when a verb is subject to the same argument requirements in simple clauses and RCs, it shows *Uniform Projection*; and thus English is classified as a language with the Uniform Projection Property (UPP), whereas Mandarin lacks this property[31]. We

[31] According to Williams (2008), Igbo, a language of southeastern Nigeria, demonstrates similar lack of UPP.

will show that this language-specific property may interact with the learning bias to generate expectations that vary from language to language.

	Simple Transitive Sentence		**Resultative Construction**	
	Structure	**Semantic Relation**	**Structure**	**Semantic Relation**
English	S – V - DO	DO is the patient of the verb, as well as the patient of the sentence	S – M – DO – R	DO is the patient of the M verb
Mandarin			S – M – R – DO	DO is the patient of the sentence, thus can be interpreted as either the patient or the instrument of the M event.

Table 5.1: Grammatical difference between English and Mandarin

5.2 Implications for Acquisition

Now consider what implications these cross-linguistic variations have for acquisition. I have emphasized, for several times in previous sections, that the distinction between the verb's event and the clause's event, hence the distinction between verb-based ONP and clause-clause ONP (see Section 5.1.2) may not be necessary, *if* all complex-predicate sentences the learner receives from the input have the two types of events pick the same set of participant. From the review of the cross-linguistic difference in Section 5.1.3, it looks like that this condition holds for English[32]: even in complex-predicate sentences where the verb's event

[32] This condition does not hold for English intransitive verbs either; for example, the sentence 'she sang her throat sore', the event of the M verb – the singing event, and the event of the clause – the

and the clause's event are distinct events, they still share participants; for example, in the events picked out by *wipe* and *wipe-clean* in sentences like (5-5a) (i.e. 'She wiped the table clean'), 'she' is the agent and 'the table' is the patient. This condition, however, does not hold for Mandarin: for example, in (5-6b), repeated here as (5-12), the event picked out by *wipe* and that picked out by *wipe-dirty* share the same agent, but have different patient – 'the table' for the former and 'the cloth' for the latter. Existence of data like these makes the distinction between verb-based ONP and clause-based ONP necessary.

(5-12) Ta ca zang le mabu.

She wipe dirty ASP cloth

She wiped something with the cloth, and the cloth turned dirty.

In what follows, I will elaborate on how verb-based ONP and clause-based ONP may guide learning in different cases (Section 5.2.1), and then discuss some options we may have in building an adequate theory of acquisition that takes into consideration cross-linguistic data from complex-predicate sentences (Section 5.2.2).

---

singing sore event, do not pick out the same set of participants. But we focus on transitive M verbs in this chapter.

5.2.1 How Verb-Based and Clause-Based ONP Guide Learning

Verb-based SBH (hence ONP) guides learning in a way that regulates the relation between verb meaning and verb valence, while clause-based SBH (hence ONP) regulates the relation between clause meaning and clause valence. Below I will elaborate how they guide learning in each of the three cases: a) the case of simple-predicate sentence; b) the case of complex-predicate sentence where the verb's event and the clause's event pick out the same set of participants; c) the case of complex-predicate sentence where the verb's event and the clause's event pick out different sets of participants.

In the case of simple-predicate sentences, clause-based ONP works in almost exactly the same way verb-based ONP does, because the event denoted by a clause is usually named by the verb; in such cases, the object NP in the sentence is informative about the meaning of the clause, which is also informative about the meaning of the verb. For example, if the learner hears a sentence 'John *gorped* the table', regardless of which version of ONP she adheres to, she builds a meaning with the logical form in (5-13), and then uses this logical form to find the event that this description satisfies; in other words, she looks in the world for an event that has 'the table' as patient (and 'John' as agent).

(5-13) [gorp(e) & agent(e, John) & patient(e, table)]

(5-14) [gorp-clean(e) & agent(e, John) & patient (e, table)]

(5-15) [gorp(e) & agent(e, John) & patient(e, cloth)]

(5-16) [gorp-dirty(e) & agent(e, John) & patient(e, cloth)]

Clause-based ONP works in a different way from verb-based ONP, however, when the event denoted by the clause is *not* named by the verb; in such cases, for child learners guided by clause-based ONP, unlike those guided by verb-based ONP, the object NP in the sentence is *only* informative about the meaning of the clause, but *not* the verb. For example, suppose a learner experiences a scenario where John uses a cloth to wipe a table and, cleaning the table but inadvertently dirty the cloth. Suppose in addition the learner hears a sentence 'John *gorped* the table clean'. If she is a verb-based ONP learner, she builds a meaning with the logical form as in (5-13), as in the case of simple-predicate sentence; under (5-13), she would take an event where the table is a patient of as a candidate meaning for the verb *gorp*. In the above scenario, the wiping event happens to be such an event. Therefore, guided by verb-based ONP, the learner may infer that *gorp* means WIPE. If she is a clause-based ONP learner, (5-14) would be the logical form she builds; under (5-14), she would take an event where the table is a patient of as a candidate meaning for the whole clause, or for the complex-predicate *gorp-clean*; but she would probably not arrive at the precise meaning of the verb *gorp*.

Suppose now, that the learner experiences the same scenario as described above, but hears a different sentence 'John *gorped* the cloth dirty'. If she is a verb-based ONP learner, she builds a meaning with the logical form in (5-15); under (5-15), she would take an event where the cloth is a patient of as a candidate meaning for the verb *gorp*. In the above scenario, such an event might be concepts like HOLD, SWING, etc., but perhaps not WIPE, because in a wiping event, the cloth is the instrument, not the patient. A clause-based ONP learner, on the other hand, builds a meaning with the logical form in (5-16); under (5-16), she would take an event where the cloth is a patient of as a possible meaning of the whole clause (or of the complex-predicate *gorp-dirty*), of which the verb *gorp* picks out a sub-event. A clause-based ONP learner, in this case, would not get decisive information about what the verb *gorp* means, but the cues she gathers would not mislead her at least; and she could then combine the cues gathered here and those gathered from other sentences where the same verb occurs (e.g. 'John *gorped* the table') to arrive at a more precise meaning of the verb *gorp*. A verb-based ONP learner, on the other hand, would form a hypothesis about the precise meaning of the verb *gorp* more quickly, but at a risk: Suppose *gorp* happens to mean WIPE, but not HOLD or SWING (as in Mandarin sentences like (5-6b), or (5-12)); then this learner would be misled by holding the wrong expectations. Worse, if she infers from 'John *gorped* the table' that *gorp* means WIPE, but infers from 'John *gorped* the cloth dirty' that *gorp* means HOLD or SWING, she would have to resolve this conflict.

The learner presumably learns from every sentence she is exposed to, simple- as well as complex-predicate ones. In face of the cross-linguistic data, therefore, for Mandarin, clause-based ONP may be safer, because for input sentences like (5-6b, 5-7b, 5-8b, 5-9b), clause-based ONP prevents the learner from forming wrong hypotheses about the verbs' meanings; for English that do not have data like those, however, verb-based ONP may be more effective to guide the learner to arrive at a precise meaning of the verb.

5.2.2 Towards an Adequate Acquisition Theory about ONP

Now, for theorizing about ONP, clearly, remaining ignorant about the distinction between the verb's event and the clause's event is not wise. With that distinction in mind, there are multiple options for building the cross-linguistic variation into the theory of ONP.

First, certain filter on the learner's intake could be imposed. Specifically, the ONP expectation could specify within itself that complex-predicate sentences are not materials for this bias to operate on. As a consequence, the learner, regardless of the target language (English-like, or Mandarin-like), would only deploy ONP on simple-predicate sentences. In such a case, the distinction between the verb's event and the clause's event, and hence verb-based ONP and clause-based ONP, does not need to enter the theory.

Second, it is possible that all learners, regardless of the target language, all entertain one version of ONP; in such a case, the theory can also hold without specifying the verb- vs. clause-based ONP distinction. However, independent accounts will be necessary to explain either of the following two cases. If it is verb-based ONP that guides all learners, learners of Mandarin and English, then, an account that explains how Mandarin-learning children eventually retreat from the wrong hypotheses made on basis of verb-based ONP would be necessary. On the other hand, if it is clause-based ONP that guides all learners, then, an account that explains how English-learning children come to know that their language does not allow sentences like '*John wiped the cloth dirty' would be necessary. These accounts are possible all external to the theory of ONP, but nevertheless requires additional explanation.

Third, it is also possible that ONP takes a general form that holds across languages, one that makes the general prediction as in (5-17), which learners of all languages respect; at the same time, learners of different languages may entertain different specific formulations of this general expectation – for example, English-learners may expect (5-18a), as predicted by verb-based ONP, whereas Mandarin-learners may expect (5-18b), as predicted by clause-based ONP. In such a case, an account that explains how learners find the appropriate specific formulation for the target language in the first place will be necessary.

(5-17) Direct objects name event patients.

(5-18a) Direct objects name patients of the main verbs' events.

(5-18b) Direct objects name patients of the sentences' events.

Each option is possible, and it requires lots of work to decide between them. But what seems to be a necessary first step is to look at young learners' use of the ONP expectation at some early developmental stage. In particular, we may examine Mandarin- and English-learners' use of ONP in guiding their expectations about the thematic relation of the direct object of RCs to its verb, to see whether their expectations are in accordance with the language-specific properties. If no, then the third option seems less likely. If yes, then bot the first and the second option may be suspended, because they clearly utilize ONP to guide their interpretations of RCs (i.e. complex-predicate sentences), and they entertain language-specific versions of ONP. Therefore, as an initial attempt to decide among the multiple possible options towards an adequate theory of ONP, we decided to take a look at Mandarin- and English-learning children's thematic interpretations of the direct object NP in resultative constructions.

5.3 Experiment Overlook

5.3.1 Research Question

ONP predicts that the learner expects the direct object of a sentence to name the patient of the event. In light of the distinction between the verb's event and the clause's event made visible by the context of complex-predicates, and in light of cross-linguistic variations between Mandarin and English in RCs, the current study asks the following question: Do children learning English and Mandarin utilize different versions of ONP to guide their interpretation about the thematic relation of the direct object to the main verb in a resultative sentence? Specifically, for the direct object of a resultative sentence, do children expect it to name the patient of the event of the means verb, or the patient of the event of the clause; and do English- and Mandarin-learners differ in their interpretations?

5.3.2 Plan

We conducted two experiments, one with adult speakers of English and Mandarin (Experiment 1, Section 5.4), the other with child learners of English and Mandarin (Experiment 2, Section 5.5). Both experiments used the same materials, but with different methods. Experiment 1 was conducted for the purpose of control - to test how well the experimental design tapped into the facts of the adults' grammars, and to establish adults' behavioral portfolio for comparison with children. Below I discuss the general design idea that applies to both experiments.

5.3.3 Experiment Design

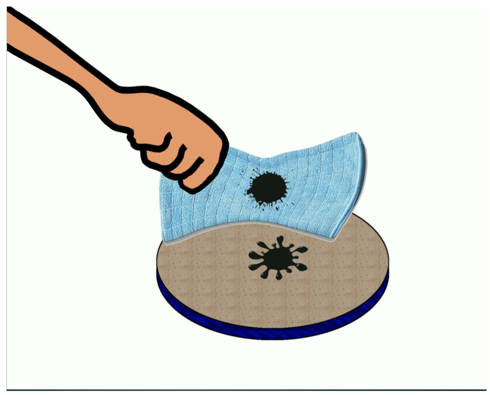
To address the research question, we designed a novel-noun reference-resolution task, using a 2 (language: English vs. Mandarin) by 3 (condition: RES, SC, CON) between-participants design. Participants were asked to determine the referent of the direct object that was denoted by a novel NP in resultative sentences (RES condition), in simple transitive sentences (SC condition), or in linguistically neutral sentences (CON condition); see (5-18) for examples.

(5-18a) RES: She wiped the *zop* clean.

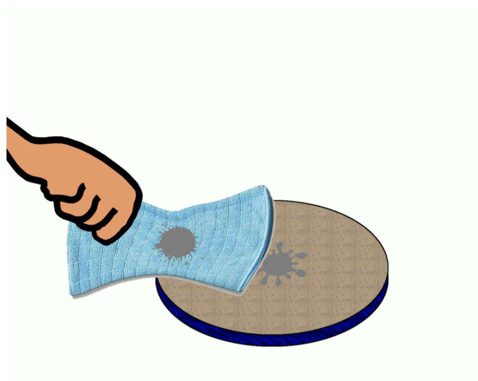
(5-18b) SC: She wiped the *zop*.

(5-18c) CON: It's a *zop*.

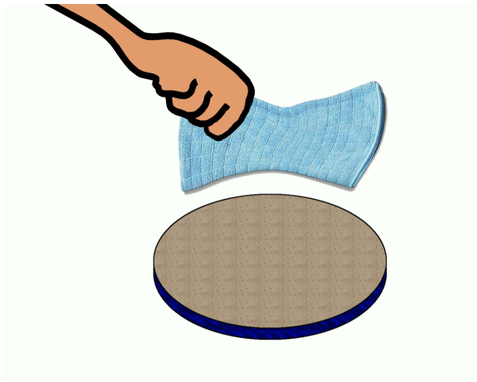
The task was designed in a way such that the referent of the novel noun was ambiguous between two objects. This ambiguity was introduced by presenting some complex causative events, the general schema of which was: a hand (i.e. representing the agent) used one object (i.e. the designated instrument) to act upon another object (i.e. the designated patient); as a result of this action, both objects entered the same result state. See Figure 5.1 for an illustration of the schema, using the wiping-clean event (and its sub-events) as an example. The labels 'the designated instrument' and 'the designated patient' are used for the purpose of clarity and simplicity. We use the word *designated* to highlight the point that the two objects were the instrument and patient of the main verb's event from the experiment designers' perspective. However, these labels may not hold in two cases, first, for the resultative sentence (e.g. 'She wiped the *zop* clean'), both objects are patients of the clause's event, by virtue of entering the same result state caused by the means verb; and second, participants in the experiments are not guaranteed to view the stimuli and assign thematic roles in the same way as we do.



(a)



(b)



(c)

Figure 5.1: Illustration of the wiping-clean event – (a) initial state; (b) intermediate state; (c) end state

In this design, the non-linguistic stimuli made available two candidates for the novel noun's referent: the object corresponding to the designated instrument, and the object corresponding to the designated patient. Meanwhile, participants may also bring in their expectations about how to interpret the linguistic stimuli, which may shape their interpretations of the novel noun. Of particular interest is child learner's choice in the resultative sentence: if they entertain verb-based ONP, expecting that the direct object names the patient of the verb's event, then the referent of the novel noun would be identified as the designated patient; if they entertain clause-based ONP, expecting that the direct object names the patient of the clause's event, then both objects could be the referent of the novel noun.

In addition to the primary research question, the design also allows room for discussions about a secondary question. As discussed extensively in Chapter 4, any stretch of world could be viewed in many different ways. The complexity of the nonlinguistic stimuli in this design makes available multiple ways to construe the event. Of salience are at least these events: the means event (e.g. wiping), the two result events – each object entering a result state (e.g. cleaning of the table, cleaning of the cloth), the means-result events (e.g. wipe-clean the table, wipe-clean the cloth). In particular, the means event might be construed, for example as a wiping event of the table - this is the most natural and canonical way of perceiving the event from the experiment designers' perspective (thus designating the table as the patient of the means the event, and the cloth the instrument); however, it is not impossible for the means event to be construed as a wiping event of the cloth, because one might imagine the hand is wiping the cloth against the table, and the design that the cloth also undergoes a change of state might make this construal more likely. Therefore, for participants in the SC condition (e.g. 'she wiped the *zop*'), after zooming in on the two-participant event category (according to PAM, 2 arguments, 2 participants), there are two possible two-participants events that both meet the description of the verb/clause; which event the child learner attends to, therefore, determines which object she thinks is the referent of the novel noun. As we will see in the experiments, while adult participants seemed to entertain the canonical construal – the same construal as us experimenter designers entertain, child participants might entertain some non-canonical construal, which also varied across languages.

5.3.4 Age of Child Participants

We look at children of 2.5 years of age.

Previous literature has documented the onset of RC production is around this age, although complete mastery is not achieved until quite late: Bowerman (1982) reported that English-learning children started to produce RCs at around age 2 (e.g. 'wipe table clean'), but their productions at this stage were restricted to acceptable combinations heard in the input; they did not begin to produce novel constructions creatively until after 3.5 years of age (e.g. 'pulled it unstapled' – 3;8; 'washing me blind' – 5;6); and utterances that would be judged ungrammatical by English-speaking adults were found from children as old as 9 years (e.g. 'the doggie bited him untied'). Chen (2006) looked at Mandarin-learning children's development of verb compounding, including RCs, and showed an early onset of RC productions (as early as 1;9), which were nonetheless restricted to utterances modeled from the input; productive use of RCs did not begin until 2;6; and mastery of subtle semantic constraints (such as the strict ordering of component verbs, the constrain against specifying more than one path in a single clause, the constraints on the second predicate that are inherently agentive; see Chen (2006), pp. 115-116) was not fully achieved as old as 8;1. Snyder and Stromswold (1997) demonstrated that English RCs are acquired together with a cohort of other constructions (including datives, verb-particle-constructions, *put*-locatives, causatives) as a group that is identified as a single syntactic class, the acquisition of which is roughly located at around 2 to 2.5 years of age.

We look at this issue with children of 2.5 years of age, which is approximately the earliest documented stage of RC knowledge, because looking at the onset might shed some light on the origin of any bootstrapping strategy.

5.4 Experiment 1 - Adults

Experiment 1 is a control experiment conducted on adult English and Mandarin adult speakers, with the goals to test how good the above-discussed design taps into the facts of adult grammars, and also to establish an understanding about how other adults (than experimenter designers) view the complex events in this experiment.

5.4.1 Participants

The group comprised 24 English-speaking adults and 24 Mandarin-speaking adults, 8 participants in each condition. English adults were college students enrolling in introductory-level courses of Linguistics Department at the University of Maryland, who participated in this experiment as part of the course requirement; and Mandarin adults were recruited from the Chinese community around campus, for whom $5 was offered as monetary reward.

5.4.2 Stimuli

The non-linguistic stimuli were 7 short videos (7 trials), in each of which two novel objects were involved in some causative event. Each video included the following parts: a) familiarization with two objects by presenting them one after another, b) a causative event where an animated human hand was using one object to act on the other, and both of the objects underwent some change of state and entered the same result state (see Figure 5.1 in Section 5.3.1 for an illustration of the causative event), and c) the two objects were presented widely apart on each side of the screen, in their end state (the state after change), and the sides where the designated patient and instrument showed up were counterbalanced across trials. All participants received the same non-linguistic stimuli.

Participants also received linguistic stimuli accompanying the non-linguistic stimuli: in part a), when participants saw presentation of the two objects, they heard 'wow, look at that' and 'do you see that'; in part b), when participants saw the causative event, they heard a resultative sentence if they were in RES condition, a simple clause if they were in the SC condition, or a linguistically-neutral sentence if they were in the CON condition; see (5-18) in Section 5.3.1 for example sentences; and in part c), when they saw the two objects presented on opposite sides of the screen, they were asked to identify the reference of the novel noun with a prompt like 'which one is the *zop*'.

There were 7 trials, each with a different causative event corresponding to two predicates (one means predicate M, one result predicate R); so there were 14 predicates appearing in the experiment in total, and they were summarized in Table 5.2. All the predicate words are known to adult participants, but child participants may only know a subset of them, and individual differences in predicate knowledge are expected.

English		**Mandarin**	
M	**R**	**M**	**R**
wipe	clean	ca	ganjing
pound	flat	da	ping
scratch	rough	gua	lan
push	crooked	ji	wan
poke	red	chuo	hong
rub	black	mo	hei
bump	awake	zhuang	xing

Table 5.2: Predicate words used in Experiment 1 and 2

All participants in all conditions saw the same set of 7 events, but heard different sets of 7 sentences: in each trial, both of the two predicates occur within a resultative frame in RES condition; only the means predicate occurs in the SC condition, within a simple clause frame; and neither of them occurs in CON condition's neutral frame. The two objects in each trial were designed to be of as equal novelty as possible such that children[33] would not easily assign the novel noun to one of the objects because she already has a name for the other. The 7 events and the corresponding 14 predicates were selected carefully based on the following criteria: a) the result event denoted by the result predicate is a likely

[33] For adults, we used the same materials as that used for children.

state for both objects to enter[34]; b) the action of the means event is picturizable and is properly described by the means verb; c) the combination of the means and result predicate respects the semantic constraints of resultative construction in each language.

5.4.3 Method

Adult participants were tested in the form of Multiple Choice Questions, where they were asked to explicitly provide their answer about the referent of the novel noun, by indicating their choice among three options: A. the object on the left, B. the object on the right, and C. either. With the 'either' option available, choosing A implies disallowance of B and vice versa, but choosing C implies allowance of both.

5.4.4 Apparatus

The linguistic stimuli were recorded by native speakers of English and Mandarin in child-directed speech[35], and were edited using the Praat program. The non-linguistic stimuli – the novel objects and the events – were made using the

[34] We understand it is very hard to make the likelihood of entering the result state equal for the designated patient and designated instrument, because it is usually expected and thus pragmatically more felicitous for the patient of the means event to undergo some change of state while maintaining the state of the instrument as original. But we tried our best to make the result state to be a likely state for the instrument to enter.

[35] We used the same materials for adults, so adults also heard child-directed speech.

animation-making software Gimp. The stimuli were presented using the QuickTime program on a Macbook Pro computer.

5.4.5 Procedure

An adult participant either came to the Cognitive Neuroscience of Language Lab on campus or met with the researcher at some low-distraction place of their convenience. The procedure began with obtaining the participant's informed consent. He/she was then directed to the testing computer, and was given a headphone, an answer sheet and a pen. The experimenter instructed the participant in the following way: 'You will be seeing 7 short videos involving two objects, and hearing some sentences describing the videos; the sentences involve some invented nouns; after each video, the two objects will be shown on opposite sides of the screen, and there will be some questions asking which of the two objects the novel noun referred to; your task is to decide the meaning of the novel noun, by circling your answer on the answer sheet, among three pre-printed options - A. the object on the left, B. the object on the right, and C. either; for each video, you will provide one answer.' There was no requirement on timing for adults to circle their answer – they could do it any time and could also change their mind during the experiment.

5.4.6 Measurement and Coding

In this experiment, each adult gave 7 answers (one for each trial). Each option was coded as 'designated patient', 'designated instrument' or 'either', and the proportion of each option out of the 7 trials was taken as the dependent variable. Thus, for each participant, there were three measures: proportion of 'designated patient' (i.e. number of trials where the participant's option corresponded to the patient object divided by 7), proportion of 'instrument' option, and proportion of 'either'.

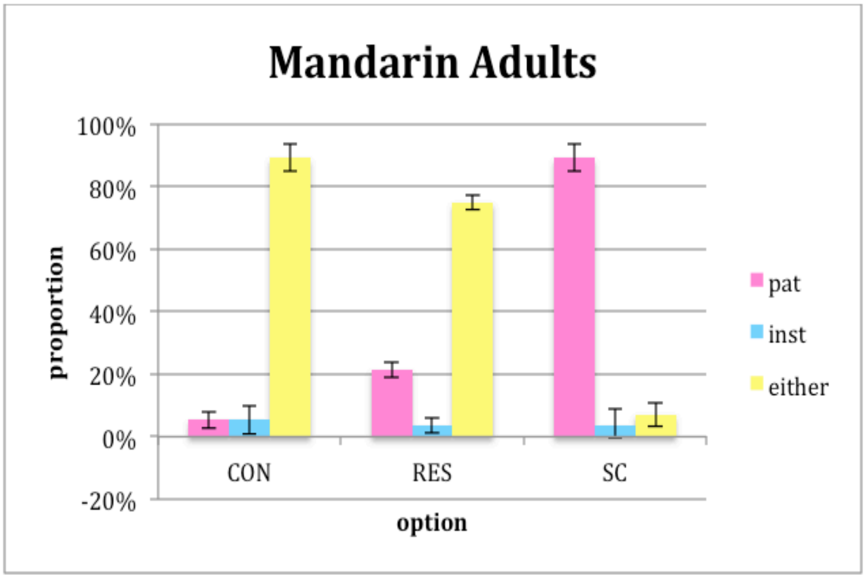
5.4.7 Predictions

If the design succeeds in tapping into the adult English and Mandarin grammars, then we would expect to see a large proportion of 'either' choice among Mandarin-speaking adults and a large proportion of 'designated patient' choice among English-speaking adults in the RES condition; and for the SC condition, we would expect to see a large proportion of 'designated patient' choice for both English- and Mandarin-speaking adults, if they view the means event in the same way as us the experiment designers do.

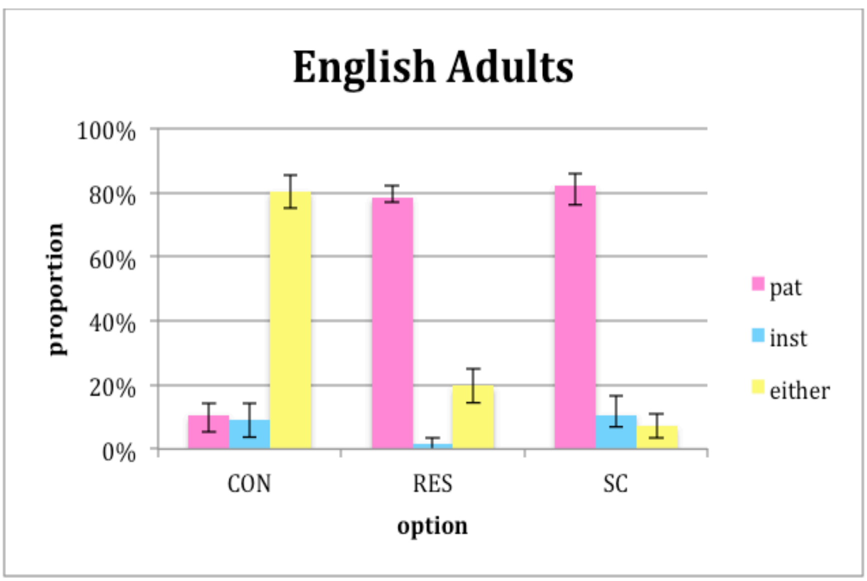
5.4.8 Results

In RES condition, Mandarin-speaking adults chose the indeterminate option 'either' (75%) significantly more than English-speaking adults did (19.6%), and also more often than the other two options as well ('patient': 21.4%, 'instrument': 3.6%); the dominant choice of English-speaking adults, in contrast was 'patient' (78.6%). This pattern is consistent with the observation that Mandarin grammar allows the direct object of a resultative sentence to be interpreted as *the patient of the clause's event*, which can sometimes be identical as the patient of the means event and sometimes as the instrument that undergoes the change of state described by the result predicate; and it is also consistent with English grammar that only allows the direct object NP to be interpreted as *the patient of the verb's event*. Interestingly, though, among the 25% of Mandarin-speaking adults' non-'either' option, 'patient' took up 21.4% whereas only 3.6% for 'instrument'. This pattern suggests a preference towards the patient object over the instrument object.

In SC condition, both Mandarin- and English-speaking adults most often chose the 'patient' option (Mandarin: 89.3%, English: 82.1%), in accordance with the grammars of both languages. In CON condition where the linguistic stimuli is not expected to exert any influence on people's interpretation, we observed exactly this expected pattern: both English- and Mandarin-speaking adults chose 'either' for most of the time (Mandarin: 89.3%, English: 80.4%), and the rest small proportion is almost equally distributed between 'patient' and 'instrument' for both languages (Mandarin – 'patient': 5.3%, 'instrument': 5.3%; English – 'patient': 10.7%, 'instrument': 8.9%). See Figure 5.2 for illustrations of the above observations.



(a)



(b)

Figure 5.2: Results with Mandarin (a) **and English** (b) **adults**

5.4.9 Discussion

The above results truthfully reflect the fact we know about the typological difference between Mandarin and English resultative constructions, proving the design a valid tool to investigate children's grammatical knowledge in resultatives. In addition, adults' choices in the SC condition suggest that the most canonical two-participant event for the simple transitive sentence is the one where the designated patient is the event patient, rather than the one where the designated instrument is the event patient, which is not surprising. Plus, adults' event construal did not seem to vary as a function of their native language. We will see whether child participants would think the same in Experiment 2.

5.5 Experiment 2 - Children

With adult control data proving the design valid, we conducted Experiment 2 on child participants.

5.5.1 Participants

Fifty-seven English-speaking children (male: 29, female: 28) with a mean age of 32;00 (range: 30;00 -34;02) and thirty-nine Mandarin-speaking children (male: 21, female: 18) with a mean age of 32;08 (range: 29;15 -34;15) participated in this study. Twenty-one additional children (10 English, 11 Mandarin) were tested but excluded from the final sample. The English group was recruited through

University of Maryland Infant and Child Studies Database, and the Mandarin group was recruited via one of the following two sources in China: a. Laboratory of Developmental Studies, Dept. of Psychology, Beijing University, and b. Kmy Baby Early Childhood Center.

5.5.2 Stimuli

We tested child participants with the same experimental materials as those used in Experiment 1.

5.5.3 Method

We adopt the Preferential Looking Paradigm (PLP), the basic idea of which is: after a short period of familiarization (the content of which varies based on the research question being asked), children are presented with two images or events on both sides of a big television screen accompanied with a test utterance, and their looking preference (i.e. looking more towards one image or event over the other) is taken as indicator of their interpretation of the test utterance. This design is based on findings that children generally look more to an image or event that matches their interpretation of the audio than to a non-matching one (Golinkoff et al., 1987; Hirsh-Pasek & Golinkoff, 1999; Spelke, 1979).

In our experiment, in the familiarization phase, child participants were presented the complex events with the two objects (i.e. the designated patient, the designated instrument) involved, concurrent with linguistic narratives containing the novel noun, the structure of which varied across conditions; and at test, children saw the two objects each on one side of the screen, with the sides where the designated patient appeared counterbalanced across trials, and children were directed to look at the screen with a prompt like ‘which one is the [novel noun]’.

5.5.4 Apparatus

For the English group, the stimuli were played on a Samsung wall-mounted 51-inch plasma television, with built-in speakers, located 66 inches away from the chair (or highchair) where the infants were seated. A Sony EVI-D100 video camera was placed directly above the TV monitor. The experimenter observes the infant from another room, and will adjust the camera (by zooming in or out) to make sure the child’s face is always centered, for the convenience of coding that will be conducted after the experiment offline. The video of the child, with a picture-in-picture display of what was on the TV screen, was captured on an iMac computer using QuickTime. For the Mandarin group, the stimuli were played using a portable version of this method, with a projector, projection screen, speakers, a computer, and a camera set up in a quiet testing room provided by the kindergarten / early education center from which the children were recruited.

5.5.5 Procedure

For the English group, the procedure began with obtaining the parent(s)' informed consent and collecting parental report of the child's productive vocabulary, which included MacArhur Communicative Development Inventory (MCDI), and the 14 predicate words particularly used in this experiment. When the child was ready, he/she was led to the test room where the TV monitor and the digital camera were located. The parent came to the test room with the infant and stayed with him/her during the entire process. The infant sat either in the parent's lap or in a highchair in front of the monitor. We took precautions to ensure that the parent could not influence the child's behavior, by explicitly instructing the parent not to direct the infant's attention in any way, and asking the parent to wear a visor (to block sight) in cases where she chose to hold the infants on her lap. Then the experimenter began the experiment in the control room next door, by setting up the computer to present the video. The whole video was about 7-minutes long. Each visit usually took half an hour, including playtime, consenting and post-experiment parent debriefing.

The procedure for the Mandarin group was roughly the same, except for the following aspects: a) the child participant was accompanied by the parent (if weekend) or the teacher of his/her class at school (if school day); b) parental report of productive vocabulary only included the 14 predicate words, because we did not have (access to) a Mandarin-version of the MCDI.

5.5.6 Coding

The videotaped films of infants' attention were digitized into a format where their eye movements could be coded on a frame-by-frame basis. A trained coder coded the film in Supercoder – a custom program for coding preferential looking videos. The sound track was removed when coding was conducted, to ensure that the coder were blind to the target-distractor positions and to condition assignment. The coder identified for each frame (30 frames per second) whether the infant's eyes were oriented to the left scene, right scene, or neither, by pressing a different key corresponding to each type of look.

5.5.7 Measurements

The dependent variable commonly used in the Preferential Looking Paradigm is the average proportion of look towards the target within a selected window. We measure the average proportion of look towards each object on the screen, i.e. the designate patient and the designated instrument, within the 2-second windows after the disambiguation point – the onset of the novel noun in each of the first prompt questions (i.e. 'which one is the [novel noun]?); see Chapter 3, Section 3.3.7 for a discussion about window selection for this method). We use the proportion of look towards the designate patient out of the total look towards either object as the dependent variable: i.e. looks to designate patient / (looks to designated patient + looks to designated instrument). For example, if within the 2-second window, that is, 60 frames, the participant looked to the patient object for

30 frames and to the instrument object for 20 frames, and was not attending to either for 10 frames, then the value of the dependent variable for this participant is going to be 30 / (30+20) = 0.6. This measure considers the relative strength of attractiveness of the two objects, the value of which will always fall in the range of [0, 1]: a value above 0.5 indicates more time spent looking towards the designated patient, a value below 0.5 indicates more time spent looking to the designated instrument, and a value near 0.5 indicates roughly the same amount of looking time.

It is well documented in the developmental literature that children sometimes reach the same developmental stage at slightly different ages and children of the same age may demonstrate different performance patterns of the same task depending on their individual levels of development. Therefore, while the above-mentioned primary measurement indexes participants' performance in the experiment, it is also important to measure individual differences. In this study, we are particularly interested in looking at how child participants' performance varies as a function of their knowledge of the particular predicates in the stimuli sentences; see Table 5.2 in Section 5.4.2 for all predicate words used in the experiment. For this purpose, we measured knowledge of the 14 predicates used in the current study for each participant, based on parental report. For the English group, with the availability of a widely-used measurement of general vocabulary size – the MacArthur Communicative Development Inventories, we also asked parents to fill out the a vocabulary list based on the Short Form Versions of the MacArthur CDI (Fenson et al., 2000). This is a measurement of children's productive (in contrast to receptive) vocabulary. And not surprisingly, we found significant correlations between general vocabulary size and knowledge of the 14 particular predicates; see Section 5.4.9 for more details.

5.5.8 Predictions

Our primary research question, is, to repeat: for the direct object of a resultative sentence, do children expect it to name the patient of the event of the means verb, or the patient of the event of the clause; and do English- and Mandarin-learners differ in their interpretations? Children's performance in the RES condition is going to shed light on this question. Predictions are: if children expect the direct object to name the patient of the means verb's event, then they would look more towards the designated patient than the designated instrument; if they expect the direct object to name the patient of the clause's event, on the other hand, they would spend equal amount of time attending to both objects, because both objects are patients from the perspective of the event of the clause.

As for the other two conditions, predictions for the CON condition are straightforward - children should attend to either object approximately equally. Predictions for the SC condition depend on how children construe the means event: if they construe the means event in the same way as experiment designers and the adult participants do - for example, as a wiping event of the table – then, we would expect them to attend to the designated patient more than the designated instrument; but if they entertain some other construal, for example, viewing the means event as a wiping event of the cloth – then, they would attend more to the designated instrument more.

In addition, given the set of predicate words used in the experiments are not all easy words, we do expect child participants to only know a subset of them; and we predict their performance in this task may vary as a function of their predicate knowledge.

5.5.9 Results

5.5.9.1 Overall analysis

A two (language) by three (condition) between-subject analysis of variance revealed no main effect of language, $F(1, 90) = 1.29$, $p = 0.26$; no main effect of condition, $F(2, 90) = 0.16$, $p = 0.85$; and no significant interaction between language and condition, $F(2, 90) = 2.51$, $p = 0.09$. See Figure 5.3 for illustration.

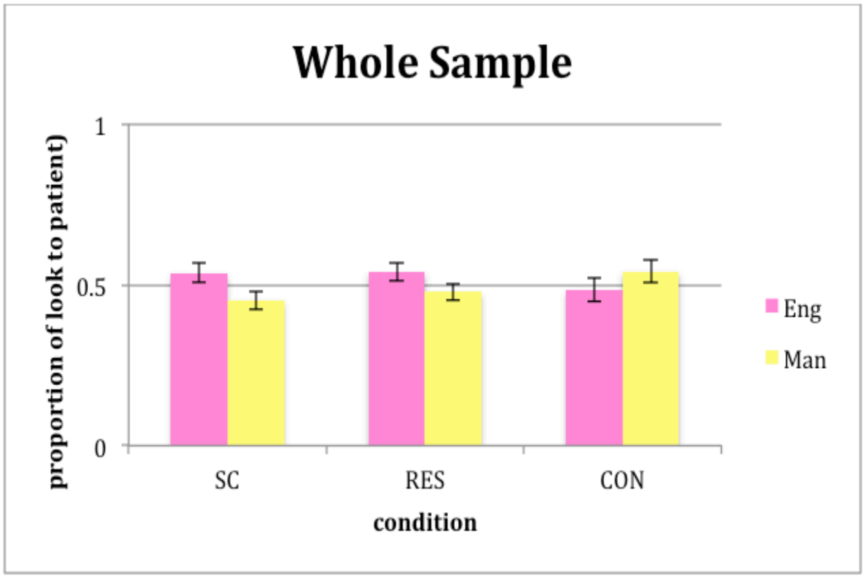


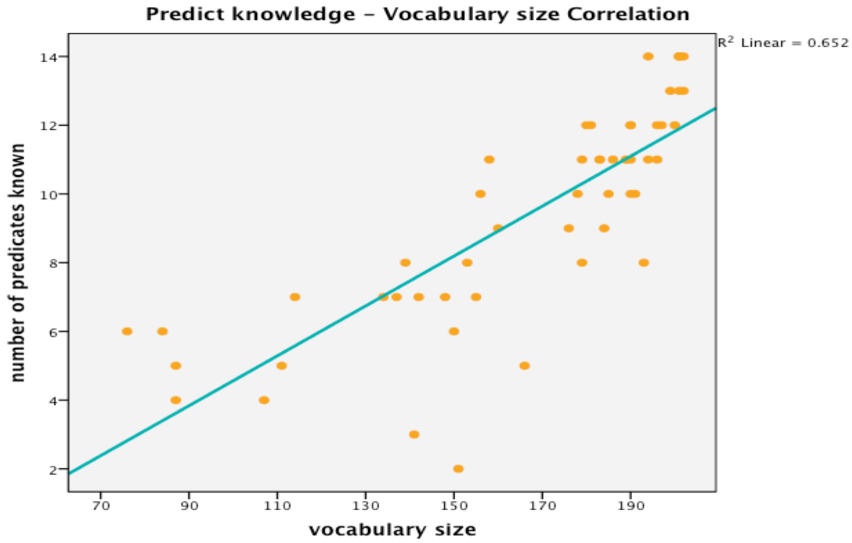
Figure 5.3: Mean look proportion towards the designated patient in three conditions for different language groups, for all participants

Figure 5.3 shows the overall pattern obtained for the whole sample, with the three conditions plotted on the x-axis, language as the grouping factor, and y-axis plots corresponding values of the dependent variable, averaged across all participants. Recall, from Section 5.4.7 that a value of the dependent variable above 0.5 indicates more time spent looking towards the patient whereas a value below 0.5 indicates more time spent looking towards the instrument.

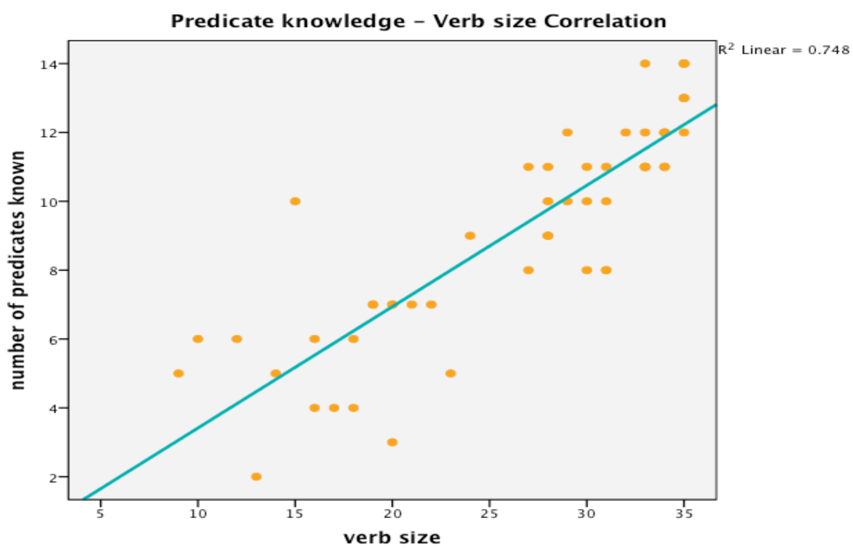
Overall, it seems children in all conditions in both language groups behaved similarly. However, with the expectation that great individual differences in predicate knowledge may exist and may hide some real effects behind the overall pattern, we took a closer look at the data.

5.5.9.2 Predicate Knowledge in Correlation with Vocabulary

A closer look at the composition of the sample reveals great individual differences in terms of vocabulary knowledge, indexed by knowledge of the particular 14 predicates in the experiment. This is not surprising, given the predicate words used in the experiment are indeed hard for this age. For the English group, predicate knowledge ranges from only knowing 2 to knowing all 14 predicates (mean = 9.13); and for the Mandarin group, it ranges from 4 to 14 (mean = 9.67). In addition, with the availability of the English group's productive vocabulary measures from MacArthur CDI, we looked at the correlation between general vocabulary size and our measure of predicate knowledge, as well as that between general verb size and predicate knowledge, and found very strong correlations, see Figure 5.4a and 5.4b. Pearson's correlation measure confirmed there was a statistically significant, strong positive relationship between vocabulary size and number of predicates known, $r(51) = 0.81$, $p < 0.01$; and there was also a statistically significant, strong positive relationship between verb size and number of predicates known, $r(53) = 0.87$, $p < 0.01$. These strong correlations between measurements of general vocabulary development and measurement of particular predicate knowledge suggest it is reasonable to use predicate knowledge as an index of individual differences of language development.



(a)



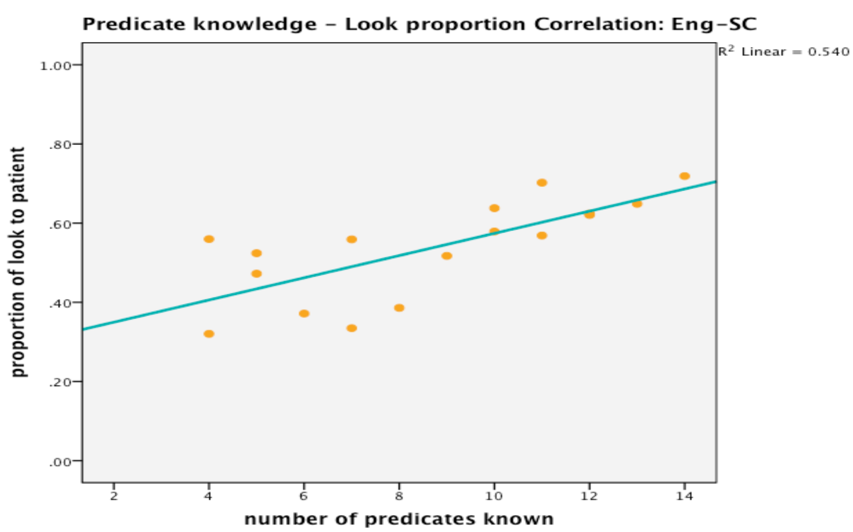
(b)

Figure 5.4: Correlation between predicate knowledge and vocabulary size (a) **and verb size** (b)

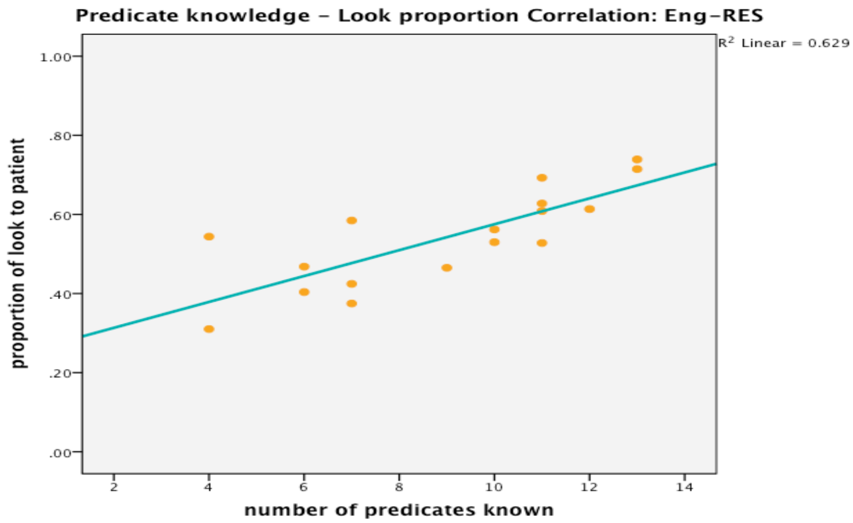
5.5.9.3 Individual Differences by Predicate Knowledge

Given the large individual differences in predicate knowledge in this sample, and given that it is frequently seen in the developmental literature that vocabulary size is often a better predictor of syntactic knowledge than age is, it is possible that some stronger effects are hidden in the overall effect, modulated by predicate knowledge.

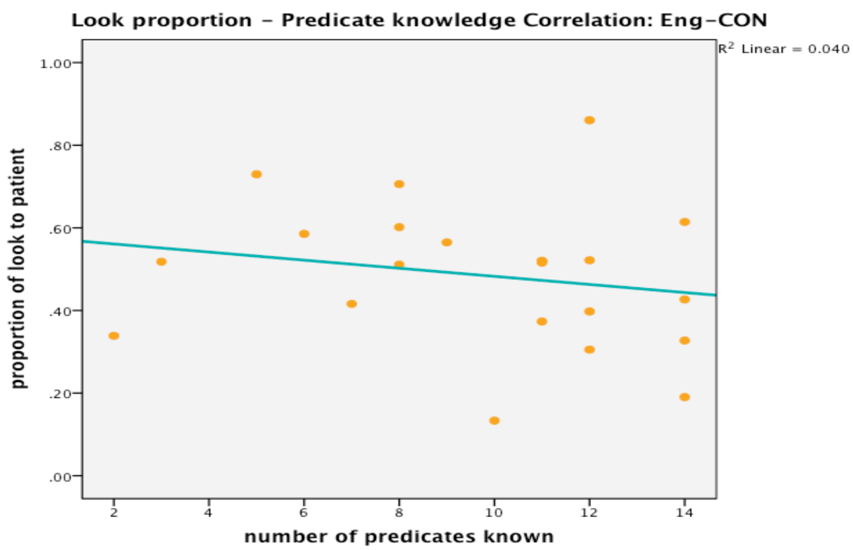
To examine this possibility, we took a look at how individual performance scores are distributed across the range of predicate knowledge. Figure 5.5 (a-f) plot the scatterplots of the relation between a participant's predicate knowledge and that participant's score of the dependent variable, in each condition, for each language group.



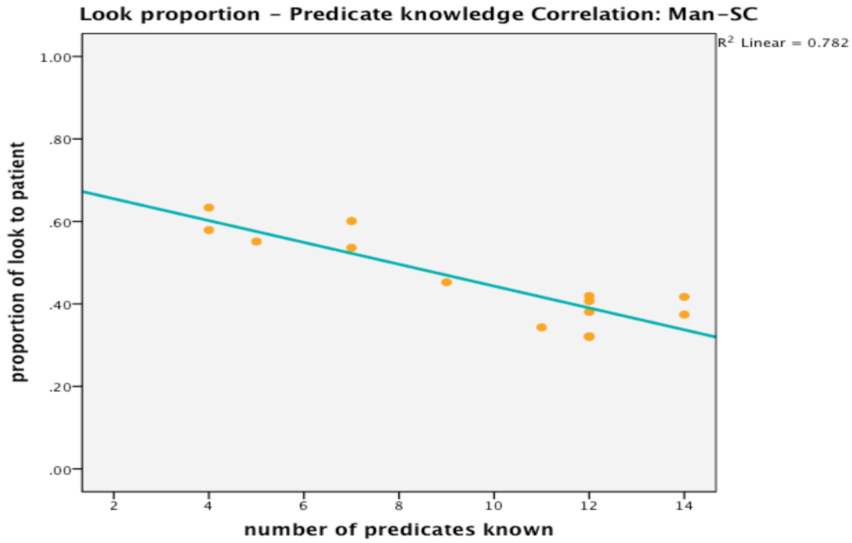
(a)


Predicate knowledge - Look proportion Correlation: Eng-RES
R^2 Linear = 0.629
proportion of look to patient
number of predicates known
1.00
.80
.60
.40
.20
.00
2
4
6
8
10
12
14

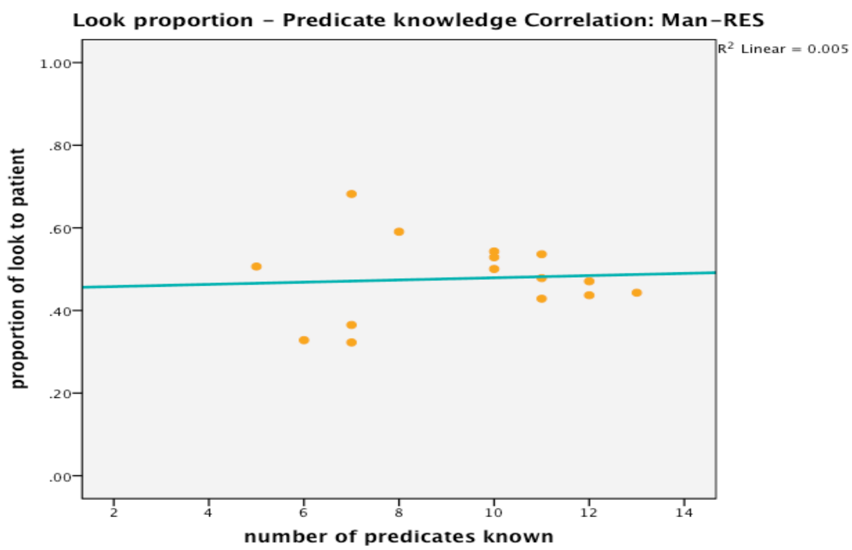
(b)


Look proportion - Predicate knowledge Correlation: Eng-CON
R^2 Linear = 0.040
proportion of look to patient
number of predicates known
1.00
.80
.60
.40
.20
.00
2
4
6
8
10
12
14

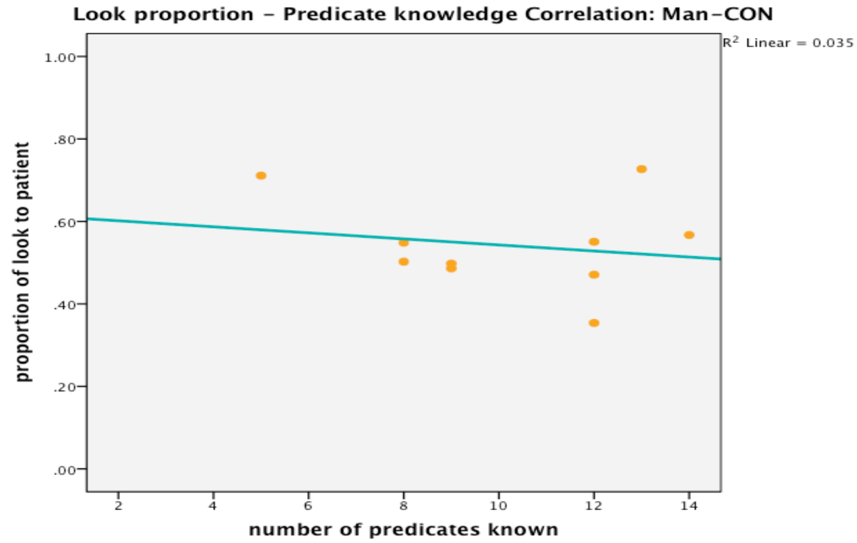
(c)


Look proportion - Predicate knowledge Correlation: Man-SC
R2 Linear = 0.782
proportion of look to patient
number of predicates known
1.00
.80
.60
.40
.20
.00
2
4
6
8
10
12
14

(d)


Look proportion - Predicate knowledge Correlation: Man-RES
R2 Linear = 0.005
proportion of look to patient
number of predicates known
1.00
.80
.60
.40
.20
.00
2
4
6
8
10
12
14

(e)



(f)

Figure 5.5: Relations between individuals' attention to the designated patient and their predicate knowledge measure (a-c: English children in SC, RES and CON conditions; d-f: Mandarin children in SC, RES and CON conditions)

From Figure 5.5, we observed the following patterns. Overall, Figure 5.5 confirms our intuition that there are huge individual differences in performance in the task, and these differences do not seem to be random, but rather, are modulated by individual predicate knowledge (at least in some conditions), which can be used as an appropriate index for individual language development. Specifically, for the English group, in SC and RES conditions, it looks like higher predicate knowledge is correlated with more time spent looking towards the patient. For the Mandarin group, in the SC condition, interestingly, higher predicate knowledge is correlated with less time spent looking towards the patient; and in the RES condition, it seems that children with higher predicate knowledge spent roughly equal amount of time looking towards either object whereas those with lower predicate knowledge's looking pattern was not regular. In contrast, in the CON condition for both language groups, no clear looking pattern was observed, nor modulated by predicate knowledge.

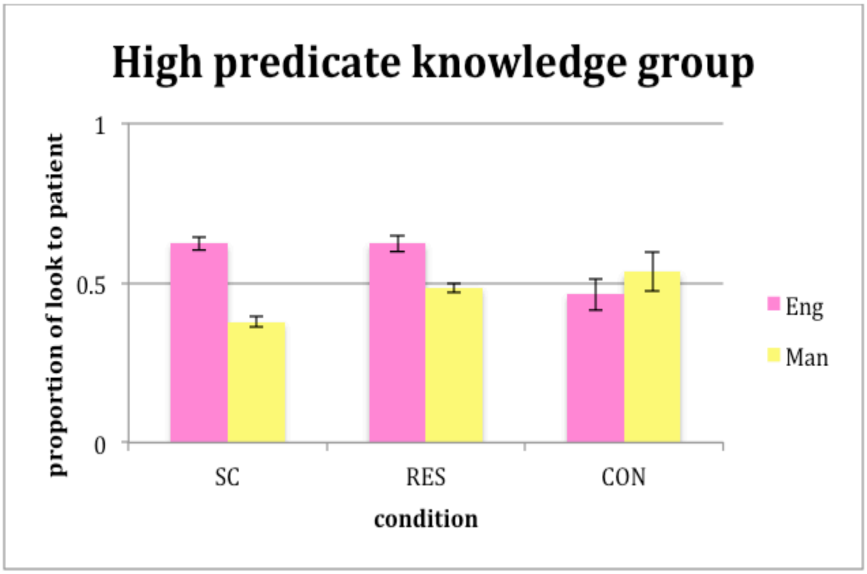
Pearson's correlation analysis shows a statistically significant, strong positive relation between the dependent variable and predicate knowledge for SC condition in the English group, $r(15) = 0.74$, $p < 0.01$; a statistically significant, strong negative relation for SC condition in the Mandarin group, $r(12) = -0.88$, $p < 0.01$; and a statistically significant, strong positive relation for RES condition in the English group, $r(15) = 0.79$, $p < 0.01$. In contrast, there is no strong correlation for CON condition in the English group, $r(20) = -0.20$, $p = 0.37$, nor in the Mandarin group, $r(8) = -0.19$, $p = 0.61$. These all confirms the above observations. However, this relation for RES in Mandarin is not strong, $r(13) = 0.07$, $p = 0.81$. But we think the lack of correlation should not be explained away by simply stating that predicate knowledge does not exert any influence. To the contrary, the scatterplot in Figure 5e shows a clear division of performance pattern between children whose predicate knowledge is higher and lower than 10; specifically, those higher than 10 seem to be close to 0.5, and those lower than 10 seem to be equally distributed above 0.5 and below 0.5. This distribution, we believe, is indicative of a predicate knowledge effect on Mandarin-learning children's performance in the RES condition.

5.5.9.4 Analysis by Predicate Knowledge Split

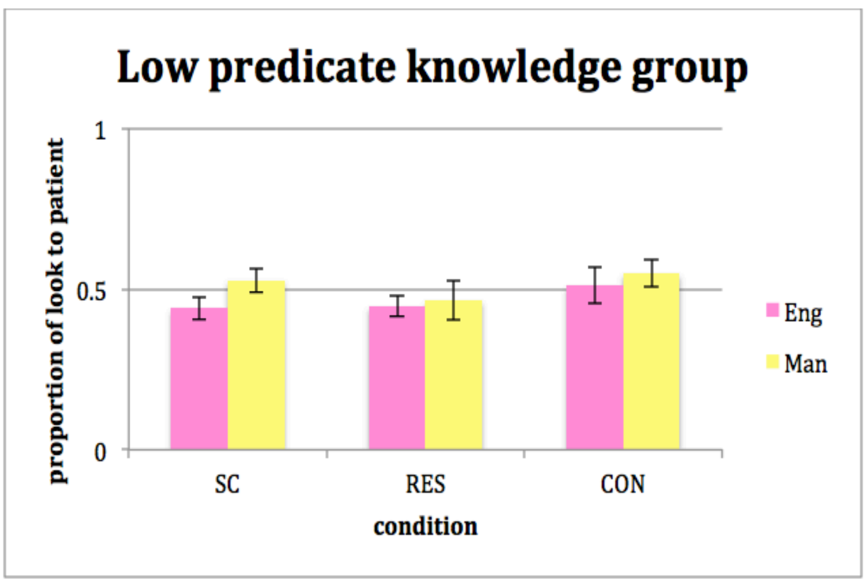
The individual differences observed in the scatterplots above (Figure 5.5) strongly suggest that the overall effect (Figure 5.3) may hide some stronger effects due to individual scores going towards different direction cancelling each other out. To examine the possible hidden effects, we split the whole sample by median predicate knowledge within each condition per language group. The median numbers of predicate words known in each condition of each language group are shown in Table 5.3. Figure 5.6 (a-b) illustrates the results of the subgroups: the group with high predicate knowledge (number of predicates known above median) in Figure 5.6a, and the group with low predicate knowledge (number of predicate known below median) in Figure 5.6b.

	English	**Mandarin**
SC	9	11.5
RES	10	10
CON	11	10.5

Table 5.3: Median number of predicate words known in each condition of each language group



(a)



(b)

Figure 5.6: Mean look proportion towards the designated patient in three conditions for different language groups, split by predicate knowledge – (a) high predicate knowledge group; (b) low predicate knowledge group

Figure 5.6 clearly shows different patterns between the two subgroups. For the group with higher predicate knowledge, basically the same pattern as in the whole sample is observed, but with a much larger effect: in both SC and RES conditions, English-learning children spent more time looking to the patient than Mandarin-learning children did; in addition, compared to the chance level (i.e. a value of 0.5), English-learning children in both SC and RES conditions looked more towards the patient than towards the instrument, while Mandarin-learning children looked more to the instrument in the SC condition but equally to the two objects in the RES condition. Both language groups in the CON condition showed a looking pattern around chance.

A 2 (language) by 3 (condition) by 2 (predicate knowledge group) independent-groups analysis of variance reveals a three-way interaction between the three factors, $F(2, 83) = 4.45$, $p = 0.02$. Since the three-way interaction is hard to interpret, we also conducted a 2 (language) by 3 (condition) independent-groups analysis of variance for each predicate knowledge group. For the high predicate knowledge group, there is a main effect of language, $F(1, 45) = 11.70$, $p < 0.01$; no main effect of condition, $F(2, 45) = 1.49$, $p = 0.24$; and an interaction between language and condition, $F(2, 45) = 8.49$, $p < 0.01$. But for the low predicate knowledge group, there is no main effect of language, $F(1, 38) = 1.50$, $p = 0.23$; no main effect of condition, $F(2, 38) = 1.14$, $p = 0.33$; and no interaction, $F(2, 38) = 0.\ 27$, $p = 0.76$.

5.5.10 Discussion

Results from this experiment suggested two things. First, children of the same age may perform in the task quite differently, modulated by their knowledge of the predicate words involved. Second, assuming it is when vocabulary barrier is minimized that children's grammatical knowledge is best revealed, then results obtained from the high predicate knowledge sub-group may be most informative to our research question.

Results from the high predicate knowledge sub-group suggested the following points. First, children in the CON condition, when hearing a linguistically neutral sentence (e.g. 'it's a *zop*') paid equal attention to the designated patient and the designated instrument, suggesting both objects were of equal interest to them; this did not vary between languages. This confirmed that the chance level (0.5 to either object) could be a valid baseline to interpret results in other conditions. Second, children in the RES condition showed different performances depending on their languages; specifically, English-learning children demonstrated a strong preference for the designated patient, whereas Mandarin-learning children paid equal attention towards both objects. This pattern was consistent with the cross-linguistic variation observed in adult grammars, namely, English requires the patient of the transitive verb to appear in the sentence whereas Mandarin allows it to be absent in a resultative sentence as along as the direct object NP refers to an object that is the patient of *the whole claus*e (and here, both objects can be the patient of the clause). Their performance in this condition, therefore, suggested that children at this age seemed to be aware of the language-specific requirements about the thematic relations of the direct object to the verb in Resultative Constructions. Third, children in the SC condition, surprisingly, also performed differently as a function of their native languages; specifically, English-learning children preferred the designated patient, whereas Mandarin-learning children preferred the designated instrument. This was surprising because in terms of grammar, English and Mandarin both dictate the object of a simple transitive sentence to be the patient of the event (verb/clause's events equivalent in this case). Although we did expect children might entertain a different way of construing the means event than adults (for example, they might see the event as a wiping event of the cloth), which might subsequently affect their choice of the patient of the means event (for example, they might think the designated instrument, the cloth, as the patient of the means event), we did not expect their construal would vary between languages. What difference between the two languages might have contributed to this? I will discuss more about this in Section 5.6.2.

5.6 General Discussion

To summarize, this chapter began with a review of existing evidence for the Syntactic Bootstrapping Hypothesis (SBH) and its universality and revealed some important limitations (Section 5.1). Current research on SBH has been limited to simple-predicate sentences, with the context of complex-predicates left uninvestigated; this context, however, is of theoretical significance, because a distinction that is not specified in the theory – the distinction between *the verb's event* (thus verb-based SBH) and *the clause's event* (thus clause-based SBH) – is only visible in this context. Moreover, current cross-linguistic investigation of SBH has mostly focused on the number-match aspect of it, but has not examined the position-role match aspect; for example, SBH predicts children expect the direct object in the sentence names the patient of the event. Taken the two limitations together, when we look at the *objects name patients* (ONP) prediction of SBH in the context of complex-predicates from a cross-linguistic perspective, a prominent question for the theory of SBH arises (Section 5.2). In particular, in Resultative Constructions (RCs) of Mandarin, the object of the sentence does not always name the patient of the main verb's event, although it does name the patient of the whole clause's event; for English, however, the object of the sentence seems to always name the patient of the main verb's event. From the learner's perspective, therefore, to be able to use the Syntactic Bootstrapping and the ONP expectation in particular as a learning guide, she has to know which event the bias is referring to – the verb's event or the clause's event, especially when the two events pick out different entities as patients, as in the case of Mandarin RCs.

In light of these cross-linguistic data, the current study conduced an experiment with Mandarin- and English-speaking adults and an experiment with 2.5-year-olds to examine their ONP expectation; in particular, how they would interpret the thematic relation of a direct object NP to the main verb in a resultative sentence. The experiments also used simple transitive sentences as a control condition. Data with adults were consistent with the cross-linguistic differences, proving our design valid. In data with children, we obtained some useful information from their performance in the RES condition to discuss our central research question and the theory of SBH, which I will discuss in Section 5.6.1; meanwhile some surprising findings from the SC condition were observed, which I will discuss in Section 5.6.2. In Section 5.6.3, I will discuss some caveats as well as some future directions.

5.6.1 Implications on the Central Question

In light of such cross-linguistic difference, I discussed three possible options for the theory of ONP (Section 5.2.2): first, impose some filter on data intake such that the learner only uses ONP with simple-predicate (but not complex-predicate) sentences; second, pick one version of ONP – verb-based or clause-based – that guides all learners, regardless of the target language, and let some other independent account to explain how learners may retreat from possible errors caused by this; and third, make available two versions of ONP – verb-based and clause-based – in the theory, for the learner select one based on their target language, leaving how such selection is made to some other account.

Much work needs to be done to decide among these options. In this current study, we took as an initial step to investigate young learners' early expectation about the thematic relation of the direct object to the main verb in RCs. This should inform us about whether young children learning different languages (Mandarin and English) hold different expectations – objects name patients of the verb's event, or objects name patients of the clause's event – at an early developmental stage. If they do have different expectations tailored to their target languages from early on, then, option three may be the best option for the theory of ONP; meanwhile, this would suggest children do use ONP on complex-predicate sentences. If they show uniform expectation across languages, then option two may be more likely; if their uniform expectation is based on the verb's event, then, what is guiding them is likely to be verb-based ONP, the version of ONP implicitly assumed in the literature; if however, their uniform expectation is based on the clause's event, then, it is probably clause-based ONP that is guiding them.

The results with a subset of the whole sample – the high predicate knowledge group – showed quite different performances between English- and Mandarin-learning children. In particular, English-learners preferred the designated patient as the referent of the novel noun – the direct object of a resultative sentence, and Mandarin-learners attended to the designated patient and the designated instrument approximately equally. These results suggested two things: first, young child learners did not seem to entertain a uniform version of ONP; second, child learners' non-uniform expectations were not random, but reflected the language-specific requirements of their target languages. These findings seem to lend more credibility to the third of the above-discussed options for the theory of ONP.

Option three, however, as already pointed out, would require an independent account to explain how young learners, at age 2.5 years, figured out the language-specific versions of ONP to entertain. This question is beyond the scope of the findings of the current experiments, and definitely requires independent work to answer. But here is one possibility. As Williams (2008) pointed out, when a verb is subject to the same argument requirements across simple- and complex-predicate structures, it show ‘uniform projection’ (See Section 5.1.3). For English, every time a transitive verb occurs, it takes its internal argument with it, no matter in a simple transitive sentence or in a complex-predicate sentence like resultatives; this is why even in complex-predicate sentence, the direct object corresponds to the patient of the event of the verb; thus English is classified as a language with the Uniform Projection Property (UPP). Mandarin, however, lacks this property, because a transitive verb may go without its internal argument in resultatives; for example, the transitive verb *wipe* must take its internal argument (e.g. the table) in simple sentences, but in resultative sentences it may ‘lose’ it (as in 5-6b, glossed as ‘she wiped cloth dirty’), and this is legitimate. Williams (2008) did not talk about UPP as a parametric variation in his paper, but it may be one, such that English is [+UPP] and Mandarin is [-UPP]. If so, then it is possible that the child learner selects which version of ONP to abide by through fixing some the specific value of this parameter for her target language.

This possibility, in turn, raises more issues. For one, it requires the learner to have fixed the value of this parameter, whatever it is, by the time SBH (hence

ONP) is used in verb learning from complex-predicate sentences, such that she is not misguided by a wrong version. It is documented in the literature that children seems to be able to use SBH as a guide in learning novel verbs from simple-predicate sentences at around 2 years of age; the parameter may not need to be set prior to this time, since both versions predict the same thing for simple-predicate clauses. But if children do use complex-predicate sentences as learning materials, they will have to fix the value of the parameter by the time the ONP bias is utilized in learning novel verbs from complex-predicate sentences; unfortunately, we do not yet have evidence about when this happens (if it does happen). A related issue concerns how the parameter of UPP is set, the challenge of which is: to know if a language has UPP requires observing for a given transitive verb, whether its patient argument always occurs with it; this does not seem to be a property that is easily observable from the input, for a learner who is still building up her verb lexicon. But remember that a parameter is a cluster of co-varying surface features explained by a single abstract representation; within the cluster, some surface facts are relatively harder to observe from the input (like the [+/-UPP] variation), some are relatively more straightforward; and the idea is, fixing the value of the easier-to-observe one(s) will automatically trigger the value(s) of those harder-to-observe ones. For example, the Compounding Parameter discussed in Snyder and Stromswold (1997) and Snyder (2001) is an abstract representation responsible for a cluster of surface facts that vary together – for instance, productive N-N compounding patterns closely with complex predicates, with the former easier to observe from the input and the latter harder, thus fixing the value of the former will trigger the value of the latter. See similar discussion about the Null Subject Parameter in Baker (2001), following (Rizzi, 1982). Following the same logic, it is possible that the [+/-UPP] variation goes together with some more observable surface variations that all subsumed by the same abstract parameter, and the value of the 'trigger' of this parameter – the observable facts – may be fixed at an early age. What exactly that cluster is, hence that trigger, is beyond the scope of current investigation.

Returning to the central concern of this chapter, results from our experiments seem to suggest that it may be necessary to include in our current theory of SBH (hence ONP) a parametric variation of two versions – verb-based and clause-based SBH (hence ONP), if the learner not only holds ONP as an expectation of interpreting thematic relations, but also uses it to learn novel verb meanings. But we should also keep in mind the caveat that our results only showed 2.5-year-olds hold different language-specific expectations, but did not show children also utilize their expectations to learn novel verbs; so, it is still possible that children use ONP to interpret thematic relations in both simple- and complex-predicate sentences, but only it to *learn verbs* from simple-predicate sentences, and thus only one ONP with regard to *verb learning* is necessary. To test this possibility, future work will need to conduct a verb learning task.

5.6.2 Surprising Finding in SC Condition

In addition to the critical condition – the RES condition, the experiments also included an SC condition as a control, where participants heard a simple transitive sentence as the linguistic stimuli (e.g. 'she wiped the *zop*'). In such sentences, because there is only one verb, the distinction between the verb's event and the clause's event does not matter, and the grammars of English and Mandarin make the same predictions about the thematic role of the direct object – the patient of the means event. The adult participants in Experiment 1 confirmed this prediction – choosing the designated patient as the reference of the novel noun.

However, given the nature of the non-linguistic stimuli used in the experiments – complex causative events that have several sub-events, we did expect there might be multiple ways to construe the means event. That was why we used the label '*designated* patient/instrument', to emphasize these were the thematic relations assigned based on the most canonical construal (according to our intuition as experiment designers), which was certainly not the only construal one could imagine. For example, in addition to viewing the means event as a wiping event of the table, it is also imaginable to view it as a wiping event of the cloth – an agent wiped the cloth (against some surface), with the cloth as the patient rather than the instrument. If this was how the means event was construed, then, the referent for the novel noun in the SC condition could be the designated instrument instead. And in fact, we did see evidence that this construal was being entertained by some children.

What was surprising, however, children's perception of the means event seemed to vary as a function of their native language. In particular, for those with high predicate knowledge, English-learning children, as expected, and as adults, preferred the designated patient as the referent of the novel noun, whereas Mandarin-learning children demonstrated an instrument preference. Here, I discuss a possible interpretation of these patterns, one that lies in the argument-drop or non-drop status of the target language.

Mandarin is an argument-drop language; English is not. Object drop in Mandarin is so common that when the object is expressed overtly by the speaker, there must be a reason. Imagine a scenario where a person uses a hammer to hit a laptop; if the laptop breaks, a Mandarin speaker would probably drop the object and say something like (5-19); but if it is the hammer that breaks, a sentence that drops the object would be pragmatically infelicitous because it fails to deliver the meaning that is normally expected; in this case, a Mandarin speaker will have to say (5-20) to draw attention AWAY from the expected meaning. Let's call this the Contrary Expectation Effect (CEE) of overt NPs in argument-drop languages.

(5-19) da sui le

hit broke ASP

Someone hit something, causing it to break.

(5-20) da sui le chuizi

hit broke ASP hammer

Someone hit something with the hammer, causing it to break.

This is consistent with the well-documented observations that zero anaphora and full NPs / overt pronouns are distributed differently: in the hierarchy of referent specificity, full NP is the most specific, followed by overt pronouns, then followed by zero anaphora; and anaphors that are more specific are used to refer to antecedents that are less accessible (Ariel, 1991; Givón, 1983; Gordon & Hendrick, 1998; inter alia). Importantly, in Mandarin, whereas for human referents there is a 3-way choice among full NP, overt pronoun (equivalent to English "she/her" or "he/him") and zero anaphora, the pronoun equivalent to English *it* is seldom overtly mentioned; so the choice for non-human objects is basically 2-way, between full NP and zero anaphora. This reduction may further magnify CEE when overt NP is used. Therefore, when Mandarin-learning children heard an overt noun in a simple transitive sentence, *and importantly*, when the accompanying scenario contains something unexpected (i.e. the change of state of the instrument), they are not only willing to, but more likely to, entertain the non-canonical event construal; in other words, it is the overt appearance of the direct object that steers children's attention away from the canonical patient to the unexpected patient (i.e. the designated instrument), resulting in more look towards the instrument object. English does not have argument-drop, therefore English-learning children naturally attend to the most canonical two-participant event and interpret the novel noun as the most 'patient-y' thing in the complex scene.

This is only our post-hoc explanation of the surprising findings. It still requires more carefully calibrated research to test this hypothesis. The findings obtained from the SC condition, together with the possible reason lying in the argument-drop status of the input language, may have the following implications. First, it resonates the issue discussed extensively in Chapter 4 – the way children view the world should not be taken for granted, but independently tested. Second, it brings up the possibility that the argument-drop status of the input language may affect children's construal of a certain scene; if it affects children's event perception, which in turn affects children's representation of the participant structure of an event, then it may also affect how children utilize SBH in real-world settings. Therefore, previous studies that did not show any particular influence of the argument-drop status of the input language on children's use of SBH (e.g. Lee & Naigles, 2008; Lidz et al., 2003) should not be taken as a conclusion for this issue; subtle differences might have been hidden, for example, by the simple event structures used in previous studies, and might be revealed when more complex events were used, as in the current experiments.

5.6.3 Some Caveats and Future Directions

In general, results from the current study are informative about the central question of concern, namely, whether it is necessary to introduce the distinction between the verb's event and the clause's event, hence two possible versions of ONP, into the theory of Syntactic Bootstrapping Hypothesis. However, there are several caveats I should point out.

First, the discussions in Section 5.6.1 on implications for the central question and in Section 5.6.2 on possible interpretation of the surprising finding are all based on results obtained from a subset of the whole sample, namely, the group of children with relatively higher predicate knowledge. Although the analysis of predicate knowledge split was well motivated, by the observation of great individual differences in performance in the task that are correlated with predicate knowledge; still, the lack of overall effect makes whatever conclusion to be drawn from the results weaker. Future experiments along the same time could take one of the following two options: a) more carefully select predicate words that child participants are likely to know; and b) conduct some kind of screening test to disqualify child participants whose predicate knowledge do not meet a pre-set criterion. Both options, however, are not easy: for a), we tried to use 'easier' words, but for the causative events used in the experiments, those words we came up with seemed to be the best choice; for b), recruiting child participants of this age is already not easy, screening will make it harder. Nevertheless, to obtain more conclusive data, such challenges need to be overcome.

Second, Mandarin-learning children's chance-level behavior in the RES condition (even for the high predicate knowledge sub-group) was the same as that in CON condition. So, it remains possible that Mandarin children simply did not do the task, but just attended to the two objects as they liked. This possibility is less likely, given they were responding to the SC condition, with a significant preference of the designated instrument. Still, perhaps they only responded to the simple transitive sentence but did not respond to the resultative sentence – treating the resultative sentence just as a linguistically neutral one, as that in CON condition. Future studies should use a design that could potentially discriminate performance in control and that in critical condition.

Chapter 6

Conclusion

This dissertation sets off with a discussion of the Poverty of Stimulus problem (Chomsky, 1980) in the realm of vocabulary acquisition, which highlights the importance of having some guidance in the word learning task. Focusing on the acquisition of verbs, this dissertation has investigated from different aspects four proposed early expectations (i.e. guides) that the learner may hold to guide her inferences about possible verb meanings, summarized in (6-1).

(6-1) Four proposed early expectations in verb learning:

1. the expectation that the grammatical category verb picks out the conceptual category event – the verb-event bias;

2. the expectation that verb meanings are relatively more general and thus broadly extendible;

3. the expectation that the number of event participants matches the number of syntactic arguments – the participant-argument-match bias (PAM), the number-match aspect of the Syntactic Bootstrapping Hypothesis (SBH);

4. The expectation that certain syntactic positions match certain thematic roles – in particular, that objects name patients (ONP), the position-role match aspect of SBH.

In four chapters (Chapter 2 – Chapter 5), this dissertation has examined each of the four expectations, hoping to complement current literature on verb learning with new data, new questions and new perspectives. In this conclusion chapter, I will first summarize key findings presented in this dissertation (Section 6.1), then discuss what these findings speak to the model of language acquisition (Section 6.2), and finally conclude the dissertation in Section 6.3.

6.1 Key Findings

6.1.1 On the Verb-Event Bias

The verb-event bias is hypothesized to narrow the learner's search space of possible verb meanings down to event concepts. Chapter 2 has examined the origin of this bias, asking whether the verb-event bias is specified within UG, or generalized inductively from experiences.

According to previous research, the verb-event bias is first seen deployed by young learners in facilitating novel verb learning not until the end of the second year (i.e. 20-24 months; Bernal et al., 2010; Oshima-Takane et al., 2011; Waxman et al., 2009), which overlaps the documented period when infants' lexicon begins to include a substantial increase of verbs (Fenson et al., 1994; Gentner, 1982; Gleitman et al., 2005; inter alia). This overlap makes it hard to decide between the two possible origins: induced based on observing many individual instances of verbs picking out events concepts, or pre-programed within UG.

The new data presented in the experiments of Chapter 2 have enabled us to pull apart the time-locked phenomena: in particular, we showed that English-learning infants were able to deploy the verb-event link in learning novel verbs at least by 18 months of age. This new finding pushes the first deployment of the verb-event bias a couple of month ahead of the 'verb spurt' period, and also pushes it a couple of month closer to the time when the linguistic and conceptual underpinnings supporting the activation of this bias are established (Gergely & Csibra, 2003; Mintz, 2006; Peterson-Hicks, 2006; Wagner & Carey, 2005; inter alia). Deployment of the bias *before* the verb spurt period makes the induction-from-experience hypothesis much less likely; and deployment of the bias not too far away from establishment of it prerequisites lends further support for the origin-from-UG hypothesis.

6.1.2 On the Specificity/Generality of Verb Meanings

The learner may expect verbs to name event concepts, but verbs vary in the degree of specificity/generality – some verbs place specific restriction on their arguments (more specific), some do not (more general). How specific/general her initial expectations of verb meanings are, remains an unanswered question. Chapter 3 investigates this question, by means of examining how the learner extends newly learned verbs to same event categories with new kinds of participants. If the learner's initial expectations of verb meanings are relatively more general, then she will be willing to extend a newly learned verb to the same event category involving a different sort of participant; if her expectations are more specific, then she may not be willing to do so.

Previous findings about this issue are mixed: some showed successful extension of novel verbs to different kinds of participants in infancy (Waxman et al., 2009), whereas some showed failure to do so even in preschool ages (Imai et al., 2005; Imai et al., 2008). I raised a hypothesis in Chapter 3 – young learners' initial expectations are relatively general, which should enable them to succeed in verb extension tasks, but they may fail to demonstrate this knowledge in tasks that overload their information processing capacities. To test this hypothesis, an experiment low in cognitive demands was designed, based on lessons learned from several previous studies. In particular, the experiment evaluates the possibility raised by Lidz et al. (2009) that the type of subject NP in the linguistic stimulus may make a difference in the information processing load imposed on infants: a lexical NP subject, by virtue of requiring processing of its lexical content, may add extra barriers for the leaner to demonstrate their knowledge, in comparison to a pronominal subject low in lexical content.

Data from this experiment have confirmed the hypotheses: first, infants of 22 months were able to successfully extend a newly learned verb to the same event category with a different sort of agent; second, their ability to do so were only seen in conditions with linguistic stimuli low in lexical content. In sum, these findings have shed light on the question of concern: verb extension ability seen with 22-month-old infants lends support for a relatively general initial expectation of verb meanings.

6.1.3 On the Participant-to-Argument Match Bias

Once the learner narrows her search space of possible verb meanings down to event concepts, under the guidance of the verb-event bias, she may further constrain her search space with the aid of cues gathered from the syntactic environment of the verb (i.e. the Syntactic Bootstrapping Hypothesis). SBH may provide a second-pass filter on what type of event a verb could mean, in one or both of the following two ways: a) number match: a sentence with n arguments pick out an event of n participants, this is called the participant-to-argument match (PAM) bias; and b) position-role match: for instance, the object of a sentence corresponds to the patient of an event; this is called the objects-name-patients (ONP) expectation.

Chapter 4 investigates an important prelude of the PAM bias – how the learner represents events and event participants. The theory of PAM states that the learner uses the number of syntactic arguments to select an event with the same number of participants as candidate event concept for the verb. Previous research (Naigles, 1990; Yuan & Fisher, 2009; Yuan et al., 2012; inter alia) testing this hypothesis has mostly followed this logic: teach child participants a novel verb embedded in an n-argument sentence (n = 1 or 2), and present them two competing scenes varying in the number of participants, one having 1 participant (as the authors presume), one having 2 (as the authors presume); if children selectively attend to the scene that has the same number of participants as arguments, then PAM is supported. This logic appears almost flawless, but is based on one important yet untested assumption: child participants would view the non-linguistic stimuli under a concept that has a particular number of participants, as experiment designers expected. Given that any stretch of the world can be viewed under many different concepts, there is no guarantee this assumption holds.

Chapter 4 puts this assumption into test, experimentally examining pre-linguistic infants' event representation. This chapter extends previous work on n = 1 or 2 to an under-investigated case – events with plausibly 3 participants, like givings, jimmyings, stealings, and beanings. Data obtained from the experiments have confirmed the intuition that adults tend to view these events under 3-participant concepts; and pre-linguistic infants too, at least with givings and jimmyings (experiments with stealings and beanings are ongoing). These results have two important implications for the theory of PAM: First, we now have a series of event types, the ways of viewing which are independently tested, such that they can be safely used in experiments testing PAM. Second, events like jimmyings (as well as stealings and beanings), like givings, are viewed under 3-participant concepts; but unlike givings, are naturally described by sentences with only 2 arguments; how PAM facilitates acquisition of these verbs, is, therefore, a serious question that the theory of PAM needs to address.

6.1.4 On the Objects Name Patients Expectation

Chapter 5 examines the position-role match aspect of SBH and focuses on the objects name patients (ONP) expectation. It asks a question about the exact nature of SBH - whether the theory refers to the event of the verb, or the event of the clause; and hence, it also asks about the nature of ONP: whether the learner expects the direct object of a sentence to name the patient of the verb's event or the patient of the clause's event.

This distinction is empty when the event of the verb and the event of the clause are identical, as in simple-predicate sentences. The two events may not be the same, however, in the context of complex-predicate sentences. Chapter 5, therefore, looks at this issue using a particular type of complex-predicate structure, the Resultative Constructions. The question is asked in light of cross-linguistic data of English and Mandarin: in English, the arguments correspond to the participants picked out by the event of the verb, whereas in Mandarin, arguments correspond to the participants picked out by the event of the clause. The cross-linguistic variation gives rise to two possible versions of ONP expectation: the verb-based ONP versus the clause-based ONP. An experiment with English- and Mandarin-learning toddlers was conducted trying to decide whether young learners entertain different language-specific versions of ONP.

Results from the experiment have suggested that 2.5-year-olds learning different languages entertained different versions of ONP in interpreting the thematic role of the direct object: English-learners expected objects to name patients of the verb (i.e. verb-based ONP), whereas Mandarin-learners expected objects to name patients of the clause (i.e. clause-based ONP). These data call for a specification of the overlooked distinction – verb's event vs. clause's event, as well as inclusion of a cross-linguistic component, in current theory of SBH.

6.2 Revisiting the Language Acquisition Model

In Chapter 1, I discussed the current investigation with respect to a language acquisition model many researchers (e.g. Lidz & Gagliardi, 2014) have been constructing. Here, in light of the new data presented in the dissertation, I will revisit this model, and discuss what information the new findings may add to the model and what questions/suggestions they bring up with respect to the model. For convenience of discussion, Figure 1.3 is copied below, as Figure 6.1.

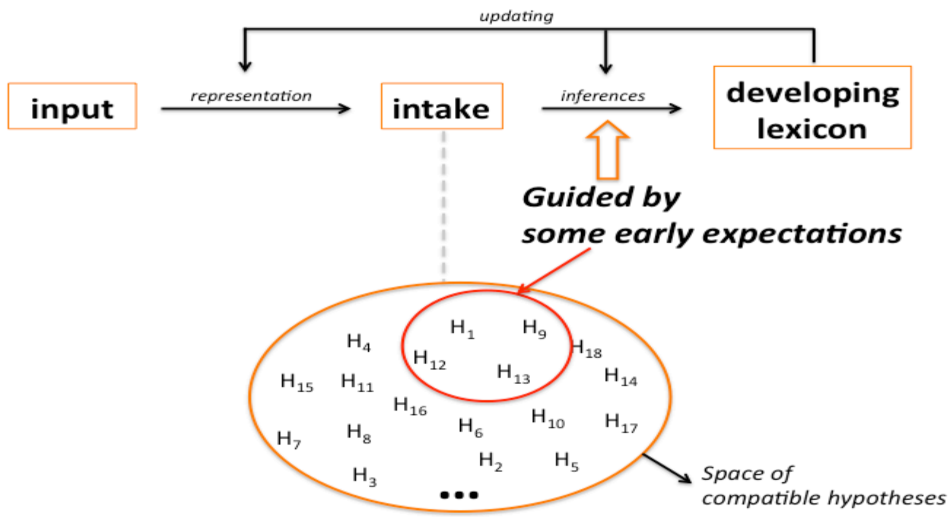


Figure 6.1 (Figure 1.3 revisited): Building a developing lexicon under guidance

To acquire a language, the learner parses the linguistic and extralinguistic input into meaningful representations that feed into latter processes – this is, to build an intake representation (arrow '*representation*'); and then, from the intake, the learner makes some 'guesses' about what the target grammar might be – this is, to make inferences about the grammar (arrow '*inferences*'). In other words, the input presents some raw linguistic phenomena, and the learner observes and represents the phenomena into some meaningful forms, and then infers what is the underlying 'machinery' that generates these linguistic phenomena. These two steps lead to a developing grammar (e.g. a developing lexicon), which in turns feeds back to the two processes with updated information (arrow '*updating*'). Importantly, both the representation and inference processes are guided by some pre-programed principles such that these processes can proceed in a constrained way. This dissertation examines how a verb lexicon is built, by specifically investigating four proposed expectations that are proposed to guide the inference process (arrow '*guided by some early expectations*'). In other words, this dissertation looks at the biases/principles/constraints that guide the learner to make 'educated guesses', rather than blind guesses, about possible meanings of new verbs.

Since the question of interest is how the learner acquires verb meanings, I will limit discussion of the model to the scope of verb learning, thus reducing each component of the model to only verb learning related information. But this simplification by no means implies verb learning is a process isolated from acquisition of other processes, and we should always keep in mind that language acquisition is a dynamic and interactive process. For instance, verb learning uses noun learning as scaffoldings, because building of an argument structure necessarily requires knowledge of nouns; for another instance, knowledge of the target language's word order is a necessary prerequisites for identifying which noun phrase is the subject and which is the object argument. With this in mind, I will proceed with some simplified illustrations of the model.

6.2.1 UG-guided Inference Mechanism

The challenge of verb learning – namely, for any given verb form, the extralinguistic context of use makes available infinitely many possible meanings – motivates an inference mechanism guided by some expectations. In Figure 6.1, the information feeding to the inference mechanism – the 'early expectations' –

could come from the UG, but could also come from the developing grammar – for example, when more and more verbs are learned, some semantic regularities could be generalized and could in turn be used as a principle guiding further learning. Data from Chapter 2 lend support for the argument that the verb-event bias is UG specified: the verb-event bias seems to lead to a growth in verb vocabulary, rather than the other way round. Data from Chapter 3 suggest that the learner seems to come to the verb-learning task expecting verb meanings are general and broadly extendible, and only acquires the specific semantic restrictions of some verbs later in development. Previous research on PAM has received support for it being an unlearned universal bias (Lidz et al., 2003). Data from Chapter 5 seem to suggest young learners, although entertain different versions, in general do respect the ONP expectation. Taken together, a simplified verb learning process guided by UG may be constructed as in Figure 6.2[36].

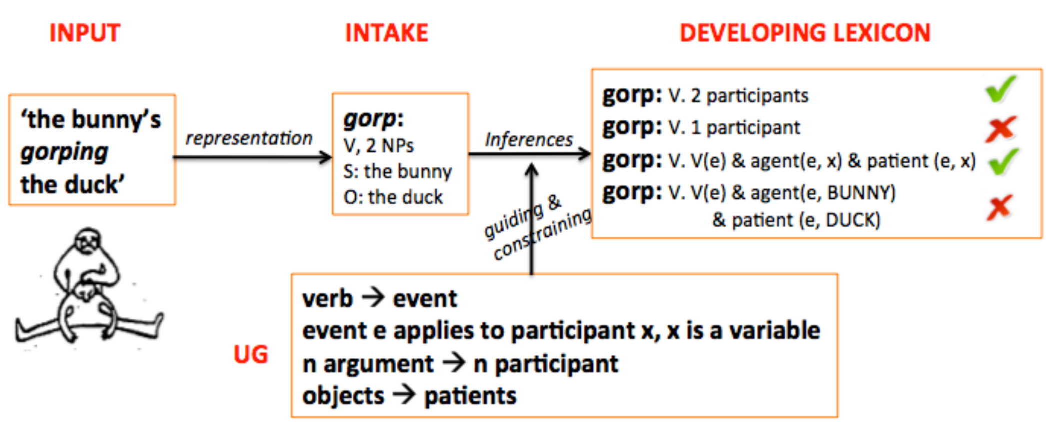


Figure 6.2: Illustration of UG-guided verb learning

[36] The picture representing the extralinguistic stimulus is taken from Naigles (1990)

From the linguistic input (e.g. ‘the bunny’s *gorping* the duck’), the learner figures out that *gorp* belongs to the grammatical category *verb* and parses the sentence into a 2-argument structure with a subject and an object; the extralinguistic context makes available infinitely many concepts to select from as candidate verb meanings. Yet, one guide from UG that constrains verb meanings to event concepts provides a first-pass filter, excluding non-event concepts; a second guide directs the learner to those 2-participant events; a third guide further zooms in on those 2-participant events whose patient is a duck; once the verb’s meaning is learner, a fourth guide tells the learner that the verb *gorp* applies to the same event category with different kinds of participants.

6.2.2 Conditions on UG Principles

Chapter 4 and Chapter 5 each points out a potential problem for Syntactic Bootstrapping Hypothesis, one for the number match bias PAM, one for the position-role match bias ONP. Chapter 4 suggests that verbs (e.g. *steal*) that can be naturally expressed in 2-argument structure (e.g. ‘Anne stole a toy’) but pick out concepts of 3 participants (e.g. thief, loot, victim) may be hard to acquire under the guidance of PAM; for example, the learner would be misled by PAM to wrongly hypothesize that *steal* means a concept like PICKUP that does not entail the third entity (i.e. the victim). Chapter 5 discusses that Mandarin-learners, if guided by the usually assumed version of ONP, verb-based ONP, would be misled by resultative sentences like ‘ta ca zang le mabu’ (glossed as: she wiped the cloth dirty), arriving at a wrong hypothesis about the meaning of *ca* ('wipe') – that it means something like SWING or HOLD (of which the cloth is a patient). What are possible solutions to these problems? One possible solution, as I mentioned briefly in the general discussion section of each chapter, might be to have the bias impose some filter on the learner's intake. Below, I discuss how this idea may or may not work.

First take for instance the problem raised by complex-predicate sentences discussed in Chapter 5. If assuming ONP only has one version that applies to all languages, and that version is the commonly assumed verb-based ONP – that is, it is *the verb's event* that the bias is referring to, then, to avoid to above-mentioned problem, ONP needs to operate on only a subset of the intake. As illustrated in Figure 6.3, ONP may exclude complex-predicate sentences from its operating scope altogether (see 'Condition'), such that the learner would only learn from simple-predicate sentences like 'she *gorped* the table' (see the arrow pointing the 'UG' box to 'inference', indicating this is an inference guided by ONP), but do not learn from complex-predicate sentences like 'she *gorped* the cloth dirty' (see the crossed-out arrow pointing from the 'UG' box to 'inference', indicating any inference made here is *not* guided by ONP). This seems to work for the problem presented in Chapter 5. However, given the data that show Mandarin-learning toddlers entertain the clause-based ONP to interpret the thematic role of the direct object in complex-predicate sentences, it seems that they do use ONP with complex-predicate sentences. But the data do not show children also utilize ONP to learn novel verbs. It is still possible that ONP is utilized in interpreting thematic roles on both simple- and complex-predicate sentences, but when it comes to learning novel verb meanings, its scope is limited to simple-predicate sentences.

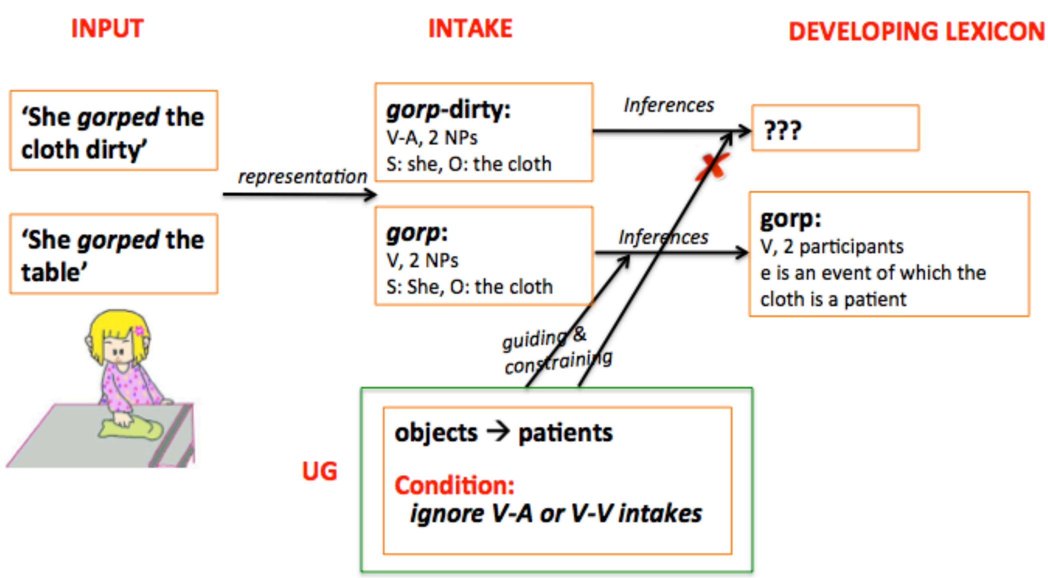


Figure 6.3: Conditions on UG principles

Now take for another instance the problem raised by verbs like *steal*, *jimmy*, and *bean* discussed in Chapter 4. The target meanings of these verbs may be hard to acquire under the guidance of PAM, and are possibly learned as PICKUP, OPEN, and HIT. Is it possible to exclude these cases by imposing a filter on the learner's intake, in a similar way as illustrated in Figure 6.3? A first thought is 'yes' – a filter that altogether excludes cases of >= 3 might work. On a second thought, however, this may not work at all. Why? Think about what 'excluding cases of >=3' means from the learner's perspective. If it means excluding sentences with more than 3 arguments, then what the learner would exclude from her learning materials are sentences like 'Anne gave betty a teddy' (3 arguments), but sentences like 'Anne stole a toy' (2 arguments) would still remain in the set of sentences she would use PAM on, which does not solve the problem at all. If it means excluding events with more than 3 participants, then the learner would exclude events that she views as having 3-participant from her hypothesis space of possible verb meanings; from the data presented in Chapter 4, we know that learners plausibly view events of jimmying, stealing and beaning as having 3-participants, but these seem to be the target concepts of the to-be-acquired verbs; therefore, excluding these events only seems to worsen the problem. Hence, filter imposing may not work for the problem presented in Chapter 4.

6.2.3 Expiration of UG principles

UG-specified expectations like syntactic bootstrapping principles (e.g. PAM, ONP), by the name 'bootstrap', are supposed to get learning off the ground; in other words, they are supposed to be guides for the first few miles of the long journey. There are many finer semantic distinctions, harder verb meanings, subtler contrasts, etc., that are beyond the scope of UG. For example, the UG-specified verb-event bias says that there are correspondences between the grammatical category *verb* and the conceptual category *event*, but such correspondences are only correlational, neither exclusive nor deterministic, since there are event concepts named by nouns (e.g. *wedding*, *earthquake*, etc.). For another example, the learner may initially expect verb meanings to be general, but there are verbs that place rather specific semantic constraints on their arguments (e.g. *eat* selects edible things as its patients, *think* select things with minds as its agents, etc.). All these need to be acquired outside the consultation of UG principles.

Some verb meanings, are not only beyond the scope of UG principles, but worse, are probably not learnable within them, like the ones discussed in Chapter 4 (e.g. *steal*, *jimmy*, *bean*). Section 6.2.2 discusses the possibility of imposing some conditions on PAM to exclude some learning materials from its operation scope, but rules it out. A less appealing idea is that these verbs are acquired after the consultation of PAM 'expires'. For example, when PAM no longer governs verb learning, sentences like 'Anne stole a toy' will no longer be used to zoom in on a 2-participant event, and *steal* will no longer be misunderstood as PICKUP. This idea necessarily admits that there might still be a period of 'misunderstanding', if PAM holds. To test this, independent tests on learners' understandings of these verbs are called for.

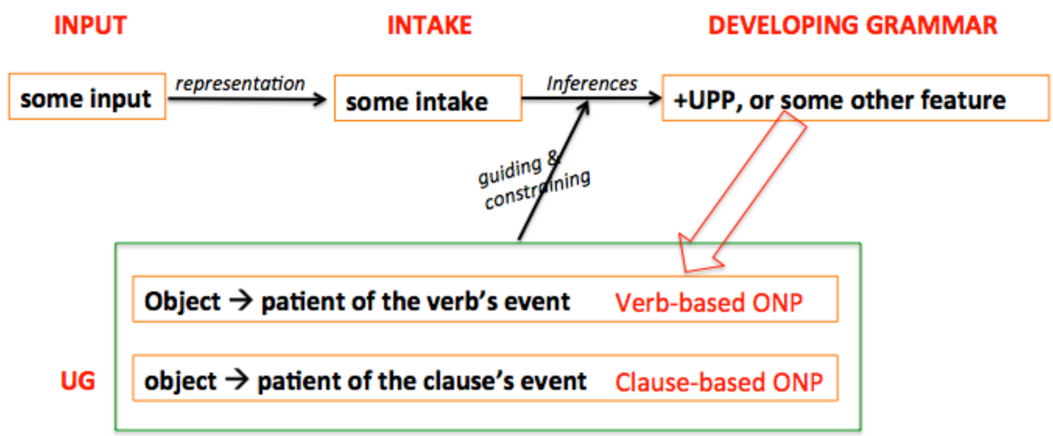
The idea of UG-principles expiration raises a whole lot of questions. For instance, how could acquisitionist theorists track the 'expiration date/period' of a UG principle? Does this date/period vary across individuals, and/or across languages? Does the expiration depend on some conditions/circumstances; and what are they? Do all bootstrapping principles expire eventually, or do they simply ‘step off the stage’ but still maintain some residual power that sometimes explains some kind of people’s ‘default’ intuition of language?

6.2.4 Selecting between Versions

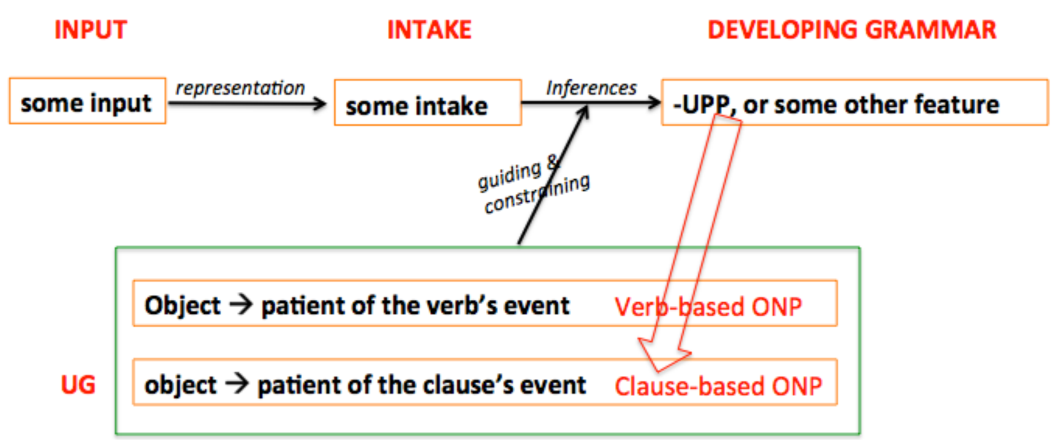
The last issue I would like to discuss is about the interaction between UG and the developing grammar. Data from Chapter 5 – that Mandarin- and English-learners are possibly entertaining two different versions of the same UG principle – highlights two points: First, UG principles are universal, but with parametric variations; for example, all learners respect the general principle ‘objects name patients’, but Mandarin-learners apply it to event of the clause, whereas English-learners apply it to event of the verb. Second, UG principles are pre-programed into the language acquisition device, but are updated with newly gained knowledge (i.e. by the developing grammar).

For example, within UG, one principle – ONP, may take two forms – verb-based ONP and clause-based ONP, selecting which depends on the ambient language. Of course the learner will not wait until figuring out the entire target grammar to make the selection, but once she fixes the value of some important parameter, she is able to make a choice. Therefore, it is something in the developing grammar that triggers the selection between verb-based ONP and clause-based ONP. That ‘something’ could be the +/-UPP feature, as discussed in Chapter 5; but the +/-UPP feature itself may not be readily observable, and may need to be triggered by some other feature of the developing grammar. Figure 6.4

(a) and (b) illustrates selection of the verb-based ONP and the clause-based ONP, respectively. It is beyond this dissertation's scope to identify what that underlying trigger is, but the idea I would like to highlight is a dynamically interactive process between UG and the developing grammar.



(a)



(b)

Figure 6.4: Selecting between verb-based ONP and clause-based ONP based on language-specific properties

6.3 Conclusion

To acquire a native language, the child learner needs to find the right 'machinery' that generates the meaningful sound waves she hears; apparently, the learner accomplishes her job rapidly, effortlessly, and with flying colors. To figure out how she accomplishes the apparently simple but indeed extremely challenging task, however, the language acquisitionists have spent years and years of hard work, gathered tons and tons of data, but are still puzzled. We are, however, happily puzzled; and delighted by every piece of new data, even though many times the new data only complicates the enigma. This dissertation sets off with the goal of contributing to solving the puzzle, but may ends up to a great extent making it more baffling. If it does, I think I can conclude by saying this dissertation contributes to the research program on child language acquisition.

BIBLIOGRAPHY

Ariel, M. (1991). The function of accessibility in a theory of grammar. *Journal of Pragmatics, 16*(5), 443-463.

Arunachalam, S., Leddon, E. M., Song, H.-J., Lee, Y., & Waxman, S. R. (2013). Doing more with less: Verb learning in Korean-acquiring 24-month-olds. *Language AcquisitionA. (2013). 20*(4), 292-304.

Arunachalam, S., & Waxman, S. R. (2011). Grammatical form and semantic context in verb learning. *Language Learning and Development, 7*(3), 169-184.

Au, T. K., & Markman, E. M. (1987). Acquiring word meanings via linguistic contrast. *Cognitive Development, 2*(3), 217-236.

Baier, R., Idsardi, W. J., & Lidz, J. (2007). *Two-month-olds are sensitive to lip rounding in dynamic and static speech events.* Paper presented at the International Conference on Auditory-VIsual Speech Processing., Groenendaal, Hilvarenbeek, The Netherlands.

Baker, M. C. (2001). *The atoms of language*. New York: Basic Books

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21*(01), 85-123.

Behrend, D. A. (1990). The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child development, 61*(3), 681-696.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, 109*(9), 3253-3258. doi: 10.1073/pnas.1113380109

Bernal, S., Dehaene-Lambertz, G., Millotte, S., & Christophe, A. (2010). Two-year-olds compute syntactic structure on-line. *Developmental Science, 13*(1), 69-76.

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development, 3*(4), 325-341.

Bowerman, M. (1982). Evaluating competing linguistic models with language acquisition data: Implications of developmental errors with causative verbs. *Quaderni di semantica, 3*(1), 5-67.

Brown, R. W. (1957). Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology, 55*(1), 1.

Bunger, A. (2006). *How We Learn to Talk About Events: Linguistic and Conceptual Constraints on Verb Learning.* (Ph.D. Dissertation), Northwestern University.

Buresh, J., Wilson-Brune, C., & Woodward, A. L. (2006). Prelinguistic action knowledge and the birth of verbs. *Action meets word*, 208-227.

Cartwright, T. A., & Brent, M. R. (1997). Syntactic categorization in early language acquisition: Formalizing the role of distributional analysis. *Cognition, 63*(2), 121-170.

Casasola, M., & Cohen, L. B. (2000). Infants' association of linguistic labels with causal actions. *Developmental Psychology, 36*(2), 155-168.

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development, 10*(2), 159-199.

Chemla, E., Mintz, T. H., Bernal, S., & Christophe, A. (2009). Categorizing words using ‘frequent frames’: what cross-linguistic analyses reveal about distributional acquisition strategies. *Developmental Science, 12*(3), 396-406.

Chen, J. (2006). The Acquisition of Verb Compounding in Mandarin. In E. V. Clark & K. B. F. (Eds.), *Constructions in acquisition* (pp. 111-135). Stanford, Calif.: CSLI Publications

Chomsky, N. (1959). A review of BF Skinner's Verbal Behavior. *Language, 35*(1), 26-58.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *The Journal of the Acoustical Society of America, 95*(3), 1570-1580.

Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). Discovering words in the continuous speech stream: the role of prosody. *Journal of phonetics, 31*(3), 585-598.

Christophe, A., Nespor, M., Guasti, M. T., & van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science, 6*(2), 211-220.

Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language, 15*(02), 317-335.

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2004). Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0). Austin: University of Texas.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28*(1), 125-127.

Dowty, D. R. (1989). On the semantic content of the notion of 'thematic role' *Properties, types and meaning* (pp. 69-129): Springer

Dowty, D. R. (1991). Thematic proto-roles and argument selection. *Language, 67*, 547-619.

Feldman, H., Goldin--Meadow, S., & Gleitman, L. (1978). Beyond Herodotus: the creation of language by isolated deaf children. In J. Locke (Ed.), *Action, gesture and symbol*. New York: Academic Press

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech, 46*(2-3), 245-264.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, i-185.

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics, 21*(1), 95-116.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. *Language Acquisition and Language Disorders, 44*, 97-135.

Fernandes, K. J., Marcus, G. F., Di Nubila, J. A., & Vouloumanos, A. (2006). From semantics to syntax and back again: Argument structure in the third year of life. *Cognition, 100*(2), B10-B20.

Fillmore, C. (1982). Frame Semantics *Linguistics in the morning calm*. Seoul: Hanshin Publishing Company

Fillmore, C. J. (1977). The case for case reopened. *Syntax and semantics, 8*(1977), 59-82.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive psychology, 31*(1), 41-81.

Fisher, C. (2002). Structural limits on verb mapping: the role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*(1), 55-64.

Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(2), 143-149.

Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive psychology, 23*(3), 331-392.

Fisher, C., Hall, D. G., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua, 92*, 333-375.

Forbes, J. N., & Poulin-Dubois, D. (1997). Representational change in young children's understanding of familiar verb meaning. *Journal of Child Language, 24*(02), 389-406.

Gagliardi, A., Bennett, E., Lidz, J., & Feldman, N. H. (2012). Children's inferences in generalizing novel nouns and adjectives *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 354-359)

Gallivan, J. (1988). Motion verb acquisition: Development of definitions. *Perceptual and Motor Skills, 66*(3), 979-986.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: language, thought, and culture* (Vol. 2, pp. 301-334). Hillsdale, NJ: Erlbaum

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve theory of rational action. *Trends in cognitive sciences, 7*(7), 287-292.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*(2), 165-193.

Gerken, L., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition, 51*(3), 237-265.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning Words and Rules Abstract Knowledge of Word Order in Early Sentence Comprehension. *Psychol Sci, 17*(8), 684-691.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*(2), 135-176.

Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study* (Vol. 3): John Benjamins Publishing

Gleitman, L. (1990). The structural sources of verb meanings. *Language acquisition, 1*(1), 3-55.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development, 1*(1), 23-64.

Gleitman, L. R., & Wanner, E. (1982). Language acquisition: the state of the state of the art In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: the state of the art* (pp. 1-50). Cambridge: Cambridge University Press

Göksun, T., Küntay, A., & Naigles, L. (2008). Turkish children use morphosyntactic bootstrapping in interpreting verb meaning. *Journal of Child Language, 35*, 291-323.

Goldberg, A. E. (1995). *Constructions : a construction grammar approach to argument structure*. Chicago: University of Chicago Press

Goldberg, A. E., & Jackendoff, R. (2004). The English resultative as a family of constructions. *Language*, 532-568.

Goldin--Meadow, S., & Mylander, C. (1984). Gestural communication in deaf children: the effects and non---effects of parental input on early language development. *Monographs of the society for research in child development, 49*(3 (207)).

Golinkoff, R., M,, Hirsh-Pasek, K., Mervis, C. B., Frawley, W. B., & Parillo, M. (1995). Lexical principles can be extended to the acquisition of verbs. *Beyond names for things: Young children's acquisition of verbs*, 185-222.

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in cognitive sciences, 12*(10), 397-403.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*(1), 23-45.

Golinkoff, R. M., Jacquet, R. C., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child development, 67*(6), 3101-3119.

Gordon, P. (2003). The origin of argument structure in infant event representations. In A. Brugos, L. Micciulla & C. E. Smith (Eds.), *Proceedings of the 28th annual Boston University Conference on Language Development*. Boston, MA: Cascadilla Press

Gordon, P. C., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive science, 22*(4), 389-424.

Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language, 51*(4), 548-567.

Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *In J. Greenberg, ed., Universals of Language. 73-113. Cambridge, MA.*

Grimshaw, J. (1981). Form, function and the language acquisition device. In C. L. Baker & J. J. McCarthy (Eds.), *The logical problem of language acquisition*. Cambridge, MA: MIT Press

Grimshaw, J. (1990). *Argument Structure*. Cambridge, MA: MIT Press

Hall, D. G., & Bélanger, J. (2005). Young children's use of range-of-reference information in word learning. *Developmental Science, 8*(1), 8-15.

Hallé, P. A., Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles? *Language and Speech, 51*(1-2), 23-44.

Hespos, S. J., & Baillargeon, R. (2001). Reasoning about containment events in very young infants. *Cognition, 78*(3), 207-245.

Hirsh-Pasek, K., Golinkoff, R., Fletcher, P., DeGaspe-Beaubien, F., & Cauley, K. (1985). *In the beginning: one-word speakers comprehend word order.* Paper presented at the Boston Child Language Conference, Boston.

Hirsh-Pasek, K., & Golinkoff, R. M. (1999). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: The MIT Press

Hirsh-Pasek, K., Golinkoff, R. M., & Naigles, L. (1996). Young children's use of syntactic frames to derive meaning. *The origins of grammar*, 123-158.

Hirsh-Pasek, K., Kemler-Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition, 26*(3), 269-286.

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy, 5*(3), 341-353.

Huang, J. (1992). Complex predicates in Control. In R. K. Larson, S. Iatridou, U. Lahiri & J. Higginbotham (Eds.), *Control and Grammar* (pp. 109–147). Dordrecht: Kluwer

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child development, 76*(2), 340-355.

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel Noun and Verb Learning in Chinese-, English-, and Japanese-Speaking Children. *Child development, 79*(4), 979-1000.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge: MA: MIT press

Jackendoff, R. (1987). The status of thematic relations in linguistic theory. *Linguistic inquiry, 18*, 369-411.

Jackendoff, R. (1990). *Semantic Structures*. Cambridge, MA: MIT Press

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & psychophysics, 14*(2), 201-211.

Johnson, E. K. (2008). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. *The Journal of the Acoustical Society of America, 123*(6), EL144-EL148.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants′ detection of the sound patterns of words in fluent speech. *Cognitive psychology, 29*(1), 1-23.

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L. J., Woodward, A. L., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive psychology, 24*(2), 252-293.

Kako, E., & Wagner, L. (2001). The semantics of syntactic structures. *Trends in cognitive sciences, 5*(3), 102-108.

Kersten, A. W., & Smith, L. B. (2002). Attention to novel objects during verb learning. *Child development, 73*(1), 93-109.

Kratzer, A. (1996). Severing the external argument from its verb. In J. Rooryck & L. Zaring (Eds.), *Phrase structure and the lexicon* (Vol. 33, pp. 109-137): Springer

Kratzer, A. (2003). *The event argument and the semantics of verbs*. Cambridge: University of Massachusetts, Amherst

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science, 218*(4577), 1138-1141.

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*: Harvard University Press

Lee, J. N., & Naigles, L. R. (2008). Mandarin learners use syntactic bootstrapping in verb acquisition. *Cognition, 106*(2), 1028-1037.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*(3), 265-288.

Li, Y. (1990). On VV compounds in Chinese. *Natural language & linguistic theory, 8*(2), 177-207.

Lidz, J. (2006). Verb learning as a probe into children's grammars. In R. M. Golinkoff & K. Hirsh-Pasek (Eds.), *Action meets word: How children learn verbs* (pp. 429-449)

Lidz, J., Bunger, A., Leddon, E., Baier, R., & Waxman, S. R. (2009). *When one cue is better than two: lexical vs. syntactic cues to verb meaning*. Unpublished manuscript.

Lidz, J., & Gagliardi, A. (2014). *How Nature Meets Nurture: Universal Grammar and Statistical Learning*. Annual Reviews of Linguistics.

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition, 87*(3), 151-178.

Lidz, J., & Gleitman, L. R. (2004). Argument structure and the child's contribution to language learning. *Trends in cognitive sciences, 8*(4), 157-161.

Locke, J. (1959). *An essay concerning human understanding*. New York: Dover Publications. Original work published in 1690

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychol Sci, 16*(8), 601-608.

Macnamara, J. T. (1982). *Names for things: A study of human learning*: Mit Press Cambridge, MA

Maguire, M. J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M., Slutzky, C. B., & Sootsman, J. (2002). *Mapping words to actions and events: How do 18-month-olds learn a verb.* Paper presented at the Proceedings of the Boston University Annual Conference on Language Development.

Marantz, A. (1984). On the nature of grammatical relations. *Linguistic Inquiry Monographs Cambridge, Mass.*(10), 1-339.

Maratsos, M., & Chalkley, M. A. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (Ed.), *Children's language* (Vol. 2, pp. 127-214). New York: Gardner

Markman, E. M. (1991). *Categorization and naming in children: Problems of induction*: Mit Press

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive psychology, 16*(1), 1-27.

Matsuo, A., Kita, S., Shinya, Y., Wood, G. C., & Naigles, L. (2012). Japanese two-year-olds use morphosyntax to learn novel verb meanings*. *Journal of Child Language, 39*(3), 637.

Mervis, C. B. (1987). Child-basic object categories and early lexical development.

Millotte, S. (2005). *Le rôle de la prosodie dans le traitement syntaxique adulte et l'acquisition de la syntax.* (unpublished PhD thesis), Ecole des Hautes Etudes en Sciences Sociales, Paris.

Millotte, S., René, A., Wales, R., & Christophe, A. (2008). Phonological phrase boundaries constrain the online syntactic analysis of spoken sentences.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(4), 874.

Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and cognitive processes, 22*(6), 898-909.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition, 90*(1), 91-117.

Mintz, T. H. (2006). Finding the verbs: distributional cues to categories available to young learners. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 31-63). New York: Oxford University Press

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive science, 26*(4), 393-424.

Morgan, J. L. (1986). *From simple input to complex grammar*: The MIT Press

Morgan, J. L., & Demuth, K. (1996). Signal to syntax: An overview. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 1-22). Mahwah, NJ: Lawrence Erlbaum Associates

Naigles, L. G. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*(02), 357-374.

Naigles, L. G., Fowler, A., & Helm, A. (1992). Developmental shifts in the construction of verb meanings. *Cognitive Development, 7*(4), 403-427.

Naigles, L. G., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child development, 64*(6), 1665-1687.

Naigles, L. R. (1998). Developmental changes in the use of structure in verb learning: Evidence from preferential looking. *Advances in infancy research, 12*, 298-318.

Naigles, L. R., & Lehrer, N. (2002). Language-general and language-specific influences on children's acquisition of argument structure: a comparison of French and English. *Journal of Child Language, 29*(03), 545-566.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht, Holland: Foris

Oshima-Takane, Y., Ariyama, J., Kobayashi, T., Katerelos, M., & Poulin-Dubois, D. (2011). Early verb learning in 20-month-old Japanese-speaking children. *Journal of Child Language, 38*(3), 455.

Peterson-Hicks, J. (2006). *The impact of function words on the processing and acquisition of syntax.* (Ph.D. Dissertation), Northwestern University, Unpublished.

Pinker, S. (1984a). *Language learnability and language development*. Cambridge, MA: Harvard University Press

Pinker, S. (1984b). The semantic bootstrapping hypothesis *Language Learnability & Language Development*. Cambridge, Mass: Harvard University Press

Pinker, S. (1989a). Learnability and cognition: The acquisition of argument structure. Cambridge, MA: MIT Press.

Pinker, S. (1989b). Resolving a learnability paradox in the acquisition of the verb lexicon. In M. L. Rice & R. L. E. Schiefelbusch (Eds.), *The teachability of language* (pp. 13-61). Baltimore, MD: Paul H. Brookes Publishing

Quine, W. V. (1960). *Word and object*. MA: MIT press

Rappaport, H. M., & Levin, B. (2001). An event structure account of English resultatives. *Language, 77*(4), 766-797.

Rappaport, M., & Levin, B. (1988). What to do with Theta-Roles in Thematic Relations. *Syntax and semantics, 21*, 7-36.

Redington, M., Crater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive science, 22*(4), 425-469.

Rizzi, L. (1982). *Issues in Italian syntax*. Dordrecht, Holland: Foris Publications

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition, 69*(2), 105-134.

Seidl, A., Hollich, G., & Jusczyk, P. W. (2003). Early Understanding of Subject and Object Wh-Questions. *Infancy, 4*(3), 423-436.

Shafer, V. L., Shucard, D. W., Shucard, J. L., & Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language, and Hearing Research, 41*(4), 874-886.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of psycholinguistic research, 25*(2), 193-247.

Shi, R. (2014). Functional Morphemes and Early Language Acquisition. *Child Development Perspectives, 8*(1), 6-11. doi: 10.1111/cdep.12052

Shi, R., Werker, J. F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy, 10*(2), 187-198.

Snyder, W. (2001). On the nature of syntactic variation: Evidence from complex predicates and complex word-formation. *Language, 77*, 324-342.

Snyder, W., & Stromswold, K. (1997). The structure and acquisition of English dative constructions. *Linguistic inquiry*, 281-317.

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition, 38*(2), 179-211.

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*(1), B1-B11.

Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology, 15*(6), 626-636.

Steele, S. (1981). *An encyclopedia of AUX: A study in cross-linguistic equivalence*. Cambridge, MA: MIT Press

Theakston, A. L., Lieven, E., Pine, J. M., & Rowland, C. F. (2002). Going, going, gone: The acquisition of the verb 'go'. *Journal of Child Language, 29*(04), 783-811.

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychol Sci, 10*(2), 172-175.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in cognitive sciences, 4*(4), 156-163.

Wagner, L., & Carey, S. (2005). 12-month-old infants represent probable endings of motion events. *Infancy, 7*(1), 73-83.

Waxman, S., & Booth, A. (2003). The origins and evolution of links between word learning and conceptual organization: new evidence from 11-month-olds. *Developmental Science, 6*(2), 128-135.

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive psychology, 43*(3), 217-242.

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive psychology, 59*(1), 67-95.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive psychology, 29*(3), 257-302.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology, 34*(6), 1289.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy, 3*(1), 1-30.

Williams, A. (2008). Patients in Igbo and Mandarin. In J. Dölling, T. Heyde-Zybatow & M. Schäfer (Eds.), *Event structures in linguistic form and interpretation* (pp. 3-30). Berlin: Walter de Gruyter

Williams, A. (2014). Causal VVs in Mandarin. In C.-T. J. Huang, A. Li & A. Simpson (Eds.), *The Handbook of Chinese Linguistics* (pp. 311-341): Wiley-Blackwell

Williams, A. (2015). *Arguments in Syntax and Semantics*. Cambridge: Cambridge University Press

Woodward, A. L., Golinkoff, R. M., Hirsh-Pasek, K., & Bloom, L. (2000). Constraining the problem space in early word learning. *Becoming a word learner: A debate on lexical acquisition*, 81-114.

Xu, F., & Tenenbaum, J. B. (2007a). Sensitivity to sampling in Bayesian word learning. *Developmental Science, 10*(3), 288-297.

Xu, F., & Tenenbaum, J. B. (2007b). Word learning as Bayesian inference. *Psychological review, 114*(2), 245.

Younger, B. A., & Cohen, L. B. (1986). Developmental change in infants' perception of correlations among attributes. *Child development*, 803-815.

Yuan, S., & Fisher, C. (2009). "Really? She Blicked the Baby?" Two-Year-Olds Learn Combinatorial Facts About Verbs by Listening. *Psychol Sci, 20*(5), 619-626.

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child development, 83*(4), 1382-1399.