

LEARNING FROM POSITIVE EVIDENCE: THE ACQUISITION OF VERB ARGUMENT

STRUCTURE

Ava Irani

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2019

Supervisor of Dissertation

Dr. Julie Anne Legate, Professor of Linguistics

Supervisor of Dissertation

Dr. Charles Yang, Professor of Linguistics

Graduate Group Chairperson

Dr. Eugene Buckley, Associate Professor of Linguistics

Dissertation Committee:

Dr. Julie Anne Legate, Professor of Linguistics

Dr. Charles Yang, Professor of Linguistics

Dr. Kathryn Schuler, Assistant Professor of Linguistics

LEARNING FROM POSITIVE EVIDENCE: THE ACQUISITION OF VERB ARGUMENT

STRUCTURE

COPYRIGHT

2019

Ava Irani

To whom it may concern.

ACKNOWLEDGMENTS

Five years ago, I embarked on a journey complete with wonderful travelling companions, winding roads, and tunnels. This dissertation is the light at the end of that PhD tunnel, and now it is time to remember those who have helped and encouraged me along the way.

I would first like to thank my advisors, Julie Anne Legate and Charles Yang, for the support, the lessons taught, and the feedback given over the last five years. Thank you both for reading and rereading many versions of this work. Your feedback was always, is always, and will always be appreciated. In particular, I would like to thank Julie for teaching me the worth of precision, and more specifically, the right way of providing evidence for new ideas by building on preexisting knowledge. The rigor in thinking that you always expected is something that I'll take along with me in life and Linguistics. Of course, if all else fails, I'll at least always have the list of syntactic tests you made us keep in Syntax II. In a similar vein, I'd like to thank Charles for teaching me to approach everything with a healthy dose of skepticism. Thanks to you, I learned to question my own ideas until I was convinced of them. I'd be remiss if I didn't also thank you for all the beers over the years— it's definitely my turn now.

Additionally, I want to thank Kathryn Schuler for her feedback on this dissertation as my committee member. Your support and collaboration on the experimental portions of this work is especially appreciated. I would also like to thank Florian Schwarz for working with me on definiteness in ASL, and for the advice on the project throughout QP1. Thanks are also due to Lauren Krawczyk for helping with recruiting and running participants, and to Stefan Pophristic for helping run participants as well. And of course, thanks to all the members of the Child Language and Learning Lab.

This dissertation has also benefited from discussions with several people over the years: Ryan Budnick (for discussions and organizing coffee hour), Aletheia Cui (for the dissertation template, proofreading QP2, and discussions on R/Python scripts in particu-

lar), Emma Nguyen, Alison Biggs, Meredith Tamminga, and all the members of Theories Being Defended (TBD). Thank you also to the members of the F-MART for their input on extremely early versions of parts of the dissertation.

There are many people I've been fortunate enough to meet during my time at Penn that I would also like to thank. Thank you Milena Šereikaitė, for being the best of best friends. I am forever grateful for your encouragement, and for the laughs we've had in this program. I would not have made it through first year without you. Thanks are also due to the other members of the class of 2019: Luke Adamson, Nattanun (Pleng) Chanchaochai, Aletheia Cui, and Kajsa Djärv, and Edward Rider-Bezerra. If it wasn't for our weekly meetings, I would be lost. Thanks, Pleng, for always being there to listen to me ramble. I will miss our Saturdays "working" together when we both lived in West Philly. Thank you, Aletheia, for always being the voice of reason. Thank you Luke for joining me to the gym each week. Thanks Kajsa, for the long conversations about everything over the years. And I would also like to thank Andrea Ceolin, for his ever-present positive vibes, and for some of the best Italian food I've ever had. I'm always happy to be your guinea pig.

Thanks are also due to Amy Goodwin Davies and Rob Wilder for their excellent advice and company throughout. Thank you Helen Jeoung for being my buddy and mentoring me through my first year in the program. Thanks are also due to Duna Gylfadóttir for showing me Philly's dance life, Jennifer Arlin for making coffee in the mornings in the Grad Lab, Yosiane White for accompanying me in *Wie is de Mol* viewings, Ruaridh Purse for fun game nights, Lacey Wade for showing me that it is possible to win at trivia, and Amy Forsyth, who in her role as administrator was always extremely helpful in dealing with any logistical matters. Others who I would like to thank are Faruk Akkus, Hezekiah Akiva Bacovcin, Sabriya Fisher, Aaron Freeman, Anton Ingason, Einar Sigurðsson, and Betsy Sneller.

Before coming to Penn, I also met several great people who've influenced my path

to Linguistics. Thank you Lori Repetti for being an extremely supportive undergraduate advisor. I am forever grateful for your encouragement as I found my way to Penn Linguistics. Thanks are also due to Mark Aronoff for mentoring my undergraduate project and providing me with excellent advice about graduate school and Linguistics. I would also like to thank Poppy Slocum, who as a graduate student herself at the time helped me apply and get accepted to my first conference. And of course, my deepest gratitude goes out to all the members of the Stony Brook Linguistics department who served as a terrific academic family to me while I was there. At Georgetown, I am indebted to Ruth Kramer, who supervised my master's thesis, and provided continual feedback on my work over the two years that I was there. I have learned many hard and soft skills from you, and I try to incorporate them in my work as often as possible.

I would also like to thank Ava Creemers, without whom this dissertation would not have been possible. Thank you for proofreading the countless drafts I sent your way, and for letting me test my ideas out on you. I learned from you that the best way to develop an analysis and improve your prose is to write them down, and edit the drafts afterwards. Thank you also for your support in the last three years that I've known you. From correcting my typos to making me coffee when I'm stressed, you have been incredible. You and your work continue to inspire me everyday.

And finally, I'd like to thank my family, who in spite of not understanding what I do, have been supportive throughout my career in Linguistics. Thanks mom and dad, for letting me chase my passion, always. The rest of my family is too big to thank individually, but my brother and grandparents deserve a special mention. Thank you all for reminding me to keep in touch; it wasn't always easy, but possible thanks to you. I love you all.

ABSTRACT

LEARNING FROM POSITIVE EVIDENCE: THE ACQUISITION OF VERB ARGUMENT STRUCTURE

Ava Irani

Supervisors: Dr. Julie Anne Legate, Dr. Charles Yang

This dissertation investigates how children acquire verb argument structure using positive evidence from the linguistic input. I discuss three case studies: the acquisition of raising and control verbs, the acquisition of causatives, and the acquisition of passives. When discussing these case studies, I address previous theories of verb argument structure learning that attempt to account for verb argument structure learning through indirect negative evidence (e.g., Pinker 1989), and I show that these approaches are inadequate in accounting for the developmental trajectory of the learner. Instead, I defend two learning models throughout the dissertation: the Sufficiency Principle (Yang 2016) and the Active Mapping Model. The Active Mapping Model of Language Acquisition illustrates how the learner uses conceptual and structural information in their language to form verb classes. In this model, there is no innate mapping between the conceptual and structural cues.

Chapter 2 discusses the acquisition of raising and control verbs. Given three kinds of verbs, verbs that are purely raising, purely control, or verbs that can be both, the literature has foregrounded a learnability problem (Becker 2006). If some verbs can be both raising and control, then what prevents the learner from assuming that all verbs can take either structure? This learnability problem, known as the problem of overgeneralization (Baker 1979), then puts forth the question of how the learner retreats from this hypothesis to arrive at the adult grammar that also has pure raising and control verbs. Using the Sufficiency Principle, I show that the problem of overgeneralization does not arise, as the number of verbs that can be both raising and control do not meet the threshold of generalization determined by the Sufficiency Principle. I also argue against theories that

propose indirect negative evidence to retreat from overgeneralization, and instead, I argue that raising and control verbs are learned from positive evidence in the input, which in this case is in the form of non-referential subjects.

In Chapter 3, I discuss the acquisition of the causative alternation rule, which is a true case of overgeneralization, as seen by the errors made by children in their production data. In the acquisition of causatives, I demonstrate that the overgeneralization errors are predicted under the Active Mapping Model where the learner categorizes verbs into classes based on conceptual and structural cues. Given the learner's vocabulary size and verb classes, the causative alternation rule is found to be productive when the input the learner receives is examined. Thus, under this learning model, the child errors are predicted. Moreover, using the Sufficiency Principle to determine the threshold of generalization, I show that the learner retreats from overgeneralization when their vocabulary size increases, as the rule is then no longer productive when the input is examined. In this chapter, I also test the Sufficiency Principle experimentally, and find support for it in the results obtained from child participants.

In Chapter 4, I use the Sufficiency Principle and the Active Mapping Model to examine the developmental trajectory of children's acquisition of passives. This chapter answers two questions: whether the passive construction is productive for the learner early on, and whether the asymmetry in the acquisition of actional and non-actional passives (e.g., Pinker et al. 1987) can be accounted for under the models of language learning assumed in this dissertation. Using the child production data and the input data, I show that the passive construction is productive in the input, and productive for the English-learning child. By examining Adam's vocabulary and verb classes, the passive rule is found to be productive for Adam given the number of passives present in the input. Moreover, under the Active Mapping Model, I show that the asymmetry in the acquisition of passives is predicted.

This dissertation investigates whether the learner's developmental trajectory can be

accounted for solely from the input data. In each of the cases discussed in the chapters of the dissertation, I demonstrate the role of the input in language acquisition. Furthermore, using the Active Mapping Model, I show that the innate linking between the verbs semantics and the syntactic structure is not only unnecessary, but also does not account for the full range of facts found in the child data.

Table of Contents

Acknowledgments	iv
Abstract	vii
List of Tables	xiii
List of Figures	xv
1 Introduction	1
1.1 Background	6
1.1.1 Theories of Verb Learning	6
1.1.2 The Problem of Overgeneralization	9
1.2 Conceptual Framework	11
1.2.1 The Threshold of Generalization	11
1.2.2 Active Mapping Model of Language Acquisition	16
2 The Acquisition of Control and Raising Verbs	21
2.1 Introduction	21
2.2 A Supposed Learnability Puzzle	26
2.3 An Indirect Negative Evidence Approach	29
2.3.1 Rate of Inanimate Subjects with Control and Raising Verbs	30
2.3.2 The Probability of Learning from Indirect Negative Evidence	34
2.4 Generalization from Positive Evidence	43
2.4.1 Positive Evidence for Learning Control and Raising Verbs	44
2.4.2 The Lack of Overgeneralization	50
2.4.3 Ambiguous vs Pure Raising Predicates	53
2.5 Conclusion	63
3 The Acquisition of Causatives	65
3.1 Introduction	65
3.2 Background	67
3.2.1 Description of Causative Errors	67

3.2.2	Experimental Evidence for the Productivity of the Causative Alternation	69
3.3	Lexical Semantic and Usage-based Approaches	73
3.3.1	Deriving Argument Structure from Verb Semantics	74
3.3.2	Usage-Based Approaches	78
3.3.3	Evaluating Preemption and Entrenchment	81
3.4	Learning Causativization	88
3.4.1	Learning Verb Classes from Structural and Conceptual Cues	90
3.4.2	Overgeneralizing The Causative Alternation	95
3.4.3	Retreat from Overgeneralization	101
3.5	Interim Summary	105
3.6	Testing the Sufficiency Principle	107
3.6.1	Background	107
3.6.2	The Present Study	110
3.6.3	Experiment 1	112
3.6.4	Experiment 2	123
3.6.5	Experiment 3	127
3.6.6	General Discussion	136
4	The Acquisition of Passives	139
4.1	Introduction	139
4.2	Background	140
4.3	Learning Passives from Positive Evidence	145
4.4	Acquisition of Passives in English	146
4.4.1	Study 1: Passives in Child-directed Input	148
4.4.2	Study 2: Productivity of Passives in Adam's Grammar	152
4.5	Acquisition of Passives in Sesotho	155
4.5.1	Background	156
4.5.2	Sesotho-Learning Children's Knowledge of Passives	157
4.6	Discussion	161
4.7	Conclusion	163
5	Conclusions	165
5.1	Summary	165
5.2	Final Remarks	167
A	Chapter 2 Appendix	170
B	Chapter 3 Appendix	171

C Chapter 4 Appendix

173

Bibliography

174

List of Tables

1.1	Summary of the number of verbs that undergo the dative alternation in the adult grammar. The dative alternation is productive for the more frequent verbs, but it is unproductive when the more infrequent verbs are taken into consideration (Yang 2016:167)	15
2.1	Inanimate subjects with raising verbs (from Becker 2014, Table 6.1)	30
2.2	Inanimate subjects with control verbs (from Becker 2014, Table 6.1)	31
2.3	Inanimate subjects with ambiguous verbs (from Becker 2014, Table 6.1) . .	32
2.4	$P(i, N, a)$ and $P(i, N, b)$ indicate the probability of a predicate being of type A with the rate of inanimate subjects at 0.2 or of type B with the rate of inanimate subjects at 0.5 at each trial. i indicates the number of inanimate subjects a predicate has occurred with. A binomial test indicated by $bt(i, N, a)$ and $bt(i, N, b)$, shows whether the number of times a predicate occurred with an inanimate subject was significant to tell whether a predicate is of type A or B. The rows where the binomial test showed significance are presented in bold. The sum of the probabilities when the binomial test showed $p < 0.05$ reveals the cumulative probabilities over the total number of trials.	36
2.5	Control, raising, and ambiguous verbs according to their frequency in the child production data in CHILDES. The threshold of generalization is calculated using the Sufficiency Principle. Evidence for the raising or ambiguous predicate rule is never sufficient enough to be generalizable. . . .	52
2.6	Non-referential subjects with <i>seem</i> and <i>begin</i> in control constructions in the input data from CHILDES. The rate of non-referential expletive subjects with <i>seem</i> and <i>begin</i> in control and raising constructions are significantly different ($p < 0.01$).	55
2.7	Non-referential subjects with <i>seem</i> and <i>begin</i> in all constructions in the input data from CHILDES. The rate of non-referential expletive subjects with <i>seem</i> and <i>begin</i> is significantly different ($p < 0.01$).	56
2.8	Inanimate subjects with <i>seem</i> and <i>begin</i> in control constructions in the input data in CHILDES. The rate of inanimate subjects with <i>seem</i> and <i>begin</i> in control and raising constructions are not significantly different ($p = 0.14$).	57
2.9	Inanimate subjects with <i>seem</i> and <i>begin</i> in all constructions in CHILDES. The rate of inanimate subjects between the verbs is not significantly different ($p = 0.24$).	58

2.10	Expletive Subjects with <i>seem</i> and <i>begin</i> in COCA. The proportion of non-referential expletive subjects with <i>seem</i> and <i>begin</i> are significantly different ($p < 0.01$).	59
3.1	Summary of verbs overgeneralized in a causative frame by the children observed by Bowerman 1982. The semantic subclass and number of errors noted with each type of verb is also provided (Pinker 1989:303, Table 7.8)	68
3.2	Raw frequency of pure intransitive verbs in child-directed speech.	88
3.3	Total number of Adam's verbs sub-divided according to the way in which the verbs were used in child-directed speech. The causative alternation rule is productive for Adam for the class of unintentional intransitives, but not for the class of intentional intransitives. The threshold of generalization was calculated using the Sufficiency Principle.	97
3.4	Total number of Ross's verbs sub-divided according to the way in which the verbs were used in child-directed speech. The causative alternation rule is productive for Ross for the class of unintentional intransitives, but not for the class of intentional intransitives. The threshold of generalizability was calculated using the Sufficiency Principle.	100
3.5	Novel items used in the artificial language. The subject <i>Vos</i> was heard and shown with each verb that was presented as a causative. Each of the five objects occurred with two of the ten verbs.	113
3.6	The distribution of verb frequency in the artificial language.	115
4.1	The 100 most frequent verbs in the CHILDES input data sorted by transitivity and intentionality.	150
4.2	The productivity of the passivization rule in child-directed input sorted by verb type. The threshold for each subclass is calculated using the Sufficiency Principle.	150
4.3	Adam's verb types and the productivity of the passive rule. The threshold for each subclass is calculated using the Sufficiency Principle.	153
4.4	Passives by verb type in the Sesotho child-direct input. The threshold of generalization is calculated using the Sufficiency Principle.	159

List of Figures

1.1	The problem of overgeneralization.	10
1.2	The Active Mapping Model of Language Acquisition does not have a built in link between the conceptual and structural cues; they are independent of each other. In this model, both conceptual and structural cues play a role in acquiring verb classes. Any rules that can be identified in the linguistic input apply at the class level.	17
2.1	Relationship between control, raising, and ambiguous verbs.	27
2.2	Number of occurrences of a predicate needed to determine whether the predicate is of type A or B.	38
2.3	Number of occurrences needed to distinguish pure control predicates from pure raising predicates with a 95% confidence interval.	40
2.4	Number of occurrences needed to distinguish ambiguous predicates from pure raising predicates with a 95% confidence interval.	41
2.5	Number of occurrences needed to distinguish pure raising predicates and ambiguous predicates with a 95% confidence interval when using the rate of inanimate subjects from <i>seem</i> and <i>begin</i> in child-directed speech.	59
3.1	Number of transitives produced by each subject in Maratsos's et al. 1987's experiment. The results of the participants illustrated here are from the non-control groups where only the intransitive form of the novel verb was used (N = 28).	71
3.2	Child C's causative errors over time (from Bowerman and Croft 2008:297). No difference was found between the rates of causative errors with verbs that do not have a suppletive form and verbs that do.	83
3.3	Child E's causative errors over time (from Bowerman and Croft 2008:297). No difference was found between the rates of causative errors with verbs that do not have a suppletive form and verbs that do.	84
3.4	The Active Mapping Model of Language Acquisition, and a summary of how verb classes are acquired. Conceptual and structural cues both play a role in acquiring verb classes, but only structural cues can apply at the class level.	89
3.5	Classification of verbs using structural and conceptual cues.	98

3.6	Summary of the learning mechanism employed by children. Rules are determined based on sufficient positive evidence in the input as determined by the Sufficiency Principle. If the rule is unproductive, the learner seeks subclasses where a productive rule may be found.	103
3.7	Summary of how the causativization rule is learned. Children first determine whether a given intransitive verb shows intentionality. Then, for each of those verbs, it is determined whether there is sufficient positive evidence to generalize the rule following the Sufficiency Principle. The causative alternation rule is generalizable when there is sufficient evidence, and it is not generalizable when there is insufficient evidence. . . .	105
3.8	A transitive causative sentence: <i>norg-ka Blerg Vos</i> in V-ka-O-S word order. Subject is shown to act on the object and slicing it into pieces with a knife. The agent is always present in the image associated with the transitive form of the verb.	114
3.9	An intransitive sentence: <i>frag Tom</i> in V-O word order. There is no agent present in the image associated with the intransitive frame.	114
3.10	Verbs used in the 5 out of 10 and 8 out of 10 language groups with their frequencies.	115
3.11	Prediction for the 5 out of 10 language group with 5 -ka verbs and 5 no -ka verbs and 8 out of 10 language group with 8 -ka verbs and 2 verbs without -ka.	118
3.12	Participants are predicted to produce -ka about a third of the time in both conditions if they probability match the token frequency of -ka in the exposure phase. The percentage of -ka production reflected in the graph is the total number of causatives produced out of the token occurrences of the 10 novel verbs.	120
3.13	Rate of -ka production across the two language groups.	121
3.14	The results for Experiment 2 show that participants overgeneralized the causative alternation rule in both the 5 out of 10 and the 8 out of 10 condition. Each participant used the causative form for each response tested.	126
3.15	Predicted results for Experiment 3. Prediction for the 5 out of 10 language group with 5 -ka verbs and 5 no -ka verbs and 8 out of 10 language group with 8 -ka verbs and 2 verbs without -ka. Participants are only expected to produce -ka for the verbs that were presented as a transitive causative. For the 8 out of 10 group, participants are expected to produce -ka across the board.	129
3.16	Summary of children's responses in Experiment 3. Children in the 5 out of 10 language group produced significantly less causatives than participants in the 8 out of 10 language group.	131

3.17	Results from testing children in Experiment 3. Number of causative and intransitive forms produced by participants grouped by the language they were tested in. All but one participant overgeneralized to some degree in the 8 out of 10 group, whereas 6 participants did not overgeneralize in the 5 out of 10 group. Almost all participants either overgeneralized completely or failed to produce to causative marker at all.	132
3.18	Results from testing children in Experiment 3. Number of causative forms produced by participants with each verb in the 5 out of 10 and 8 out of 10 language groups. The items on the x-axis are ordered according to their frequencies, with higher frequency items on the left. No significant effects of frequency were found in the production of causatives.	133
3.19	Results from testing children in Experiment 3. Number of causative forms produced by participants with each verb presented as an intransitive in the 5 out of 10 and 8 out of 10 language groups. The items on the x-axis are ordered according to their frequencies, with higher frequency items on the left. No significant effects of frequency were found between the intransitive verbs in either group.	134
4.1	The acquisition of passives using the Active Mapping Model of acquisition. Conceptual and structural cues, in this case, intentionality and transitivity, both play a role in forming verb classes. The productivity of the passive rule is then evaluated at the level of the verb class.	146
4.2	Classification of verbs using structural and conceptual cues.	149
4.3	Adam's knowledge of passives.	155
4.4	Number of passives in the input arranged by frequency.	160

Chapter 1

Introduction

Verb argument structure is a complex phenomenon, and therefore, how children come to acquire it has been the subject of much work in language acquisition studies (Goldberg, Casenhiser, and Sethuraman 2004; Gropen, Pinker, Hollander, and Goldberg 1991; Pinker 1989; Theakston 2004; Theakston, Lieven, Pine, and Rowland 2001; Wonnacott, Newport, and Tanenhaus 2008; i.a.). The question of how children learn verbs in their language remains relevant today. In this dissertation, I illustrate how verb argument structure is acquired from positive evidence in the linguistic input.

One reason for the attention this topic has received is the varied nature of the syntactic structure each verb may appear in. While many verbs can occur in multiple syntactic frames, a good portion of verbs are restricted in terms of their argument structure. For instance, some verbs can only occur as transitives (1)-(2), while others can only occur in an intransitive frame (3)-(4).

(1) *hit*

a. John hit Bill. [Transitive]

b. *John hit. [Intransitive]

(2) *touch*

a. John touched Bill. [Transitive]

b. *John touched. [Intransitive]

- (3) *vanish*
- | | |
|-------------------------|----------------|
| a. Bill vanished. | [Intransitive] |
| b. *John vanished Bill. | [Transitive] |
- (4) *fall*
- | | |
|---------------------|----------------|
| a. Bill fell. | [Intransitive] |
| b. *John fell Bill. | [Transitive] |

Moreover, the verb semantics do not form a tight correspondence with the verb's possible argument structures. Famously, *eat* and *devour* have similar meanings, but *eat* can be optionally intransitive (e.g., *I ate*), while *devour* is strictly transitive (e.g., **I devoured*). This example shows that the syntactic properties of a verb can vary, even within a single semantic class.

In addition to variation among the syntactic frames a verb can occur in, the acquisition of verb argument structure is further complicated by the fact that children use language productively. They do not simply restrict themselves to what they hear in the input, as is evident from the errors found in child speech. Some of the examples of child errors examined in this dissertation are given below.

- (5) a. I want to disappear it (3;3, Ross, MacWhinney corpus)
- b. are you going to stay me at my new school at Pittsburgh (3;5, Ross, MacWhinney corpus)
- c. if you can go it fast the pictures might run (4;7, Adam, Brown corpus)
- d. hey we're gonna fall it (4;09, Gabe, Hall corpus)

The errors found in (5) are examples of children overgeneralizing the transitive causative form to verbs that can only occur as intransitives. These examples illustrate that children form productive rules in their language, and therefore, any theory of language learning must address the question of when and how these rules are generalizable to forms that the

learner has not encountered in the input. We cannot, crucially, adopt a strict lexicalist approach (e.g., Baker 1979; Fodor 1985) to the problem, as novel verbs entering the grammar are known to adhere to a previously established pattern.

In this dissertation, I examine how children acquire verb argument structure, with a large focus on how children acquire verbs and verbal constructions that do not introduce an external argument; i.e., the argument of a verb that is introduced in its specifier. These cases are fundamental in understanding the way in which children learn verbs and verb argument structure, and they are indispensable components of research on language learning. Moreover, verbs and verbal constructions without an external argument can sometimes appear to be identical to verbs that do introduce an external argument. For instance, unaccusative and unergative verbs in English are largely indistinguishable in the structural environment they occur in. An example of this is seen when they both occur as simple intransitives as in (6).

- (6) a. I fell. [Unaccusative]
b. I ate. [Unergative]

Especially in these cases, it is important to understand the cues that are available to the child in learning verb argument structure. Research on verb argument structure has also produced a diverse range of theories of language acquisition, from strong nativist approaches to indirect negative evidence-based accounts. Through the case of the acquisition of verb argument structure, I address the different theories of language learning, and evaluate them.

In the dissertation, I discuss three instances where the learner must make decisions regarding the verb argument structure: the acquisition of raising and control constructions, the acquisition of causatives, and the acquisition of passives. Each of these cases deals with verbs or verbal constructions that do not introduce an external argument, and describes the process through which the learner can distinguish between cases where an external argument is introduced, and cases where it is not. For instance, raising verbs

such as *seem* do not introduce an external argument, but control verbs like *want* do. In learning raising and control constructions, the learner must identify which verbs are control, and which ones fall into the raising verb category. This is another instance where verbs that introduce an external argument and verbs that do not can appear in identical surface environments, as seen in (7) below.

- (7) a. John wants to like syntax. [Control]
b. John seems to like syntax. [Raising]

Second, for the acquisition of causatives, we have already seen that not all intransitive verbs can occur in transitive causative frames (3)–(4). The child must learn that some intransitive verbs do not causativize, while still maintaining that the causative alternation rule is productive in the adult grammar, since adult speakers of English willingly causativize novel intransitive verbs (Maratsos, Gudeman, Gerard-Ngo, and DeHart 1987). The case of the acquisition of causatives also directly addresses the question of how the learner identifies the verbs that introduce an external argument and verbs that do not, as the causative alternation primarily applies to the class of unaccusative verbs.

The third case, the acquisition of passives, is similarly relevant. Verbs in passive constructions do not project an external argument, and the thematic object raises to become the surface subject. This raises interesting questions for the acquisition of syntactic structure given theories like the maturation hypothesis (Borer and Wexler 1992), which proposes that A-chains are acquired late. For each of these case studies outlined above, I discuss the developmental trajectory of the child in learning verb semantics and argument structure, including any overgeneralizations that the learner makes.

In the dissertation, I also examine the problem of overgeneralization in the aforementioned cases. Children are productive learners, and therefore, in learning the rules of their language, they make overgeneralization errors, as seen in (5). This leads to the problem of overgeneralization, which refers to the idea that when the learner has generalized to a superset grammar, they cannot retreat to the subset grammar in the absence of direct neg-

ative evidence. This learnability problem is often referred to in the literature as Baker's paradox (Baker 1979). I address this problem by showing that the generalizability of any rule can be determined from the input. When there is sufficient motivation for a rule, the learner generalizes, and they retreat from their generalizations when the statistical properties of the input and the learner's vocabulary change such that there is no longer sufficient motivation. Adopting Yang's 2005, 2016 Sufficiency Principle, I show exactly when it is possible for the learner to generalize from the input.

In describing the acquisition of control and raising predicates in Chapter 2, I show that the problem of overgeneralization does not arise, even though it has been argued otherwise in the literature (e.g., Becker 2006). I present an analysis of the acquisition of control and raising predicates that illustrates the conditions under which one would expect the learner to overgeneralize, and I show that these conditions do not hold in the acquisition of raising and control verbs. In Chapter 3, I then take a look at the acquisition of causatives, which is a true case of overgeneralization. Here, I show that overgeneralization occurs in early stages of vocabulary acquisition due to the large number of unaccusative verbs in the learner's vocabulary that undergo the causative alternation in the input. Retreat from overgeneralization occurs when the learner acquires more unaccusative verbs, and the proportion of verbs that undergo the alternation is lower than the threshold of the Sufficiency Principle. Finally, I examine the acquisition of passives in Chapter 4, and determine when the passive construction is acquired productively by the learner given their vocabulary size and the data in the input.

Throughout the dissertation, I also focus on the role of the input, and how input is used in the acquisition process. In examining children's competencies of certain constructions, the data that are available to them in the input are of utmost importance. I argue that the mastery of certain constructions depends on the evidence available in the primary linguistic data, and therefore, my claims addressing the acquisition of verb argument structure clarify whether learning is driven by the data in the input, or whether it

is innately available to the learner. Here, I argue that verb argument structure is learned entirely from the learner's linguistic experience.

One aspect of language acquisition that this dissertation aims at capturing is that of individual variation between language learners. Previous approaches to the acquisition of verb argument structure (e.g., Pinker 1989) could not adequately capture the fact that the language acquisition trajectory is not the same for each child. Some children robustly produce errors while others do not. For instance, Maratsos et al. 1987 finds that many children productively generalize the causative alternation, while some do not extend the argument structure of a verb beyond what was heard in the input. The tension between proposing a general learning mechanism that also accounts for variation in the developmental trajectory is especially apparent under models that propose innate mapping between the semantics and the syntactic structure. In this dissertation, I elaborate on the individual differences between children along the way, and account for them using the Sufficiency Principle model of generalization, which is described in more detail in Section 1.2.1.

1.1. Background

The two main themes addressed in this dissertation are the acquisition of verb argument structure and the problem of overgeneralization (Baker 1979). In the subsections below, I provide an overview of previous theories on how children acquire verb argument structure. Then, I briefly discuss the problem of overgeneralization that arises in the acquisition of rules associated with verb argument structure learning.

1.1.1. Theories of Verb Learning

This section briefly describes the previous approaches taken to account for the acquisition of verb argument structure. In particular, I discuss innate syntactic and semantic

bootstrapping approaches (Fisher 2002; Fisher, Gertner, Scott, and Yuan 2010; Gleitman 1990; Pinker 1984, 1987; i.a.) and indirect negative evidence approaches (Ambridge, Pine, Rowland, and Young 2008; Bowerman and Croft 2008; Goldberg 1995; Pinker 1989; i.a.).

The mechanisms of word and verb learning proposed in the literature vary in several respects. One key point of divergence stems from the kind of equipment the learner is initially endowed with. The notion of syntactic bootstrapping, for instance, plays on the idea that the structure in which a word occurs allows the child to construe an appropriate meaning for the word (Fisher, Gertner, Scott, and Yuan 2010; Fisher, Gleitman, and Gleitman 1991; Gleitman 1990; Talmy 1975). The ability of the learner to comprehend structural relationships leads to the acquisition of word meanings. Under these approaches, children learn verbs along with their argument structure early on and use them appropriately (Borer and Wexler 1987; Crain and Lillo-Martin 1999; Gleitman and Newport 1995; Naigles 1990). This early awareness is argued to be due to the innate knowledge of the relationship between syntactic structures and the meanings they are associated with (e.g., Lidz, Gleitman, and Gleitman 2003).

The type of account outlined above does not predict child overgeneralization errors as in (5). For instance, when Ross produces *I want to disappear it* at age 3;3, it is because he assumes that *disappear* can also occur in a transitive causative form similar to a verb like *break* or *melt*. *Disappear* occurs many times in the child production data, and thus, it is clear that the learner is aware of the verb semantics and some of its argument structure properties as well. An innate linking rules approach predicts that because the learner has acquired the verb *disappear*, the relevant argument structure properties should be in place as well. If the class in which *disappear* occurs is the same class in which *break* occurs, then either both or none of the verbs are predicted to undergo the causative alternation. If the learner treats a verb like *disappear* like *break* and *melt*, by using it as a transitive causative, then it is difficult to explain how the learner retreats from this treatment of *disappear* under this approach. As a result, child overgeneralization errors are considered

to be simple lexical errors that do not result from a productive rule, otherwise, under this account, the learner would have no way of retreating from this overgeneralization.

Another nativist approach that faces the same problem with the child overgeneralizations described above is that of semantic bootstrapping (Pinker 1989; Wexler and Culicover 1980). The semantic bootstrapping hypothesis also argues for the innate linking between the syntactic structure and the semantics, similar to the syntactic bootstrapping hypothesis. The difference between the two approaches is the source of the cues available to the learner in the input. A syntactic bootstrapping approach privileges syntactic cues that allow the learner to attain the verb meaning, while a semantic bootstrapping account privileges semantic cues, where the verb argument structure is derived from the verb's meaning. Throughout the dissertation, I discuss bootstrapping hypotheses, and propose an alternative learning model in Section 1.2.2 where there is no innate linking between the verb semantics and structure.

At the other end, some learning models do not expect the learner to be able to identify and learn the full range of word meanings due to the learner's lack of understanding of the linguistic structure (Golinkoff, Shuff-Bailey, Olguin, and Ruan 1995; Hollich et al. 2000; Maguire, Hirsh-Pasek, and Golinkoff 2006). These models claim that verbs learned early on are only learned and used in specific ways. The meanings of the verb learned early on are claimed to not extend to a large range of subjects and objects associated with the verb. Models proposing the idea that child language acquisition is conservative argue that children do not extend their observations from one word to another (Brooks and Tomasello 1999; MacWhinney 2004; Theakston, Lieven, Pine, and Rowland 2001; Tomasello 2000). In learning verbs, the claim in this case is that children observing the argument structure for one verb would not extend a similar structure to other verbs. Such theories fall under the umbrella of usage-based approaches.

One aspect of a usage-based theory of acquisition (Ambridge, Pine, Rowland, and Young 2008; Goldberg, Casenhiser, and Sethuraman 2004; Tomasello 2000; i.a.) is the

argument that the learner uses indirect negative evidence in the linguistic input when acquiring verb argument structure. These analyses argue that children use statistical preemption or entrenchment-based learning when acquiring the syntactic structures associated with a verb. Under statistical preemption approaches, another more frequent form in the input that is similar in meaning rules out the necessarily infrequent ungrammatical form that is hypothesized by the learner (e.g., Bowerman and Croft 2008). For instance, the learner may have a transitive form of the verb *come* in their grammar, as in **Can you come it for me*. The transitive form of an intransitive verb like *come* is ungrammatical, and therefore, infrequent in usage. The learner might then eventually replace this form with the transitive form of the verb *bring*, which occupies a similar meaning space and is frequent in the input. Under this analysis, there crucially exists another form with a similar meaning in the input. In contrast, entrenchment approaches argue that children learn the possible syntactic frames a verb can occur in through repetition in the input (e.g., Ambridge et al. 2008). For instance, let us assume that the learner assumes that a particular verb can take an indirect object both as a prepositional phrase or in a double object construction. The more times that the verb occurs in a direct object construction, the more likely it is for the learner to assume that the indirect object cannot occur in a prepositional phrase. Throughout this dissertation, I show why the two approaches of statistical preemption and entrenchment do not adequately account for the acquisition of verb argument structure.

1.1.2. The Problem of Overgeneralization

Language acquisition research over the past several decades has also focused on what is known as the problem of overgeneralization (Baker 1979; Berwick 1985; Bowerman 1982; Pinker 1989; Yang 2016; i.a.). In learning the rules of their language, children at times make overarching generalizations resulting in a grammar that is a superset of the adult grammar (Berwick 1985). The following problem then arises: how do children unlearn

these generalizations in the absence of direct negative evidence in the input? This classic subset problem in question is represented in the diagram below:

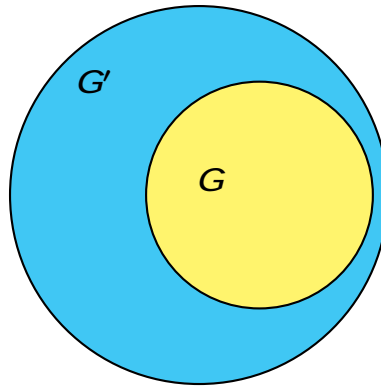


Figure 1.1: The problem of overgeneralization.

Figure 1.1 illustrates the problem of retreating from the superset grammar that the learner has acquired to the more conservative adult grammar. If the learner has acquired the superset grammar G' , the language acquisition process must be such that the learner eventually arrives at the subset grammar G . An example of the problem can be found in the case of the dative alternation (Baker 1979; Pinker 1989; Yang 2016), as mentioned above. Some ditransitive English verbs, but not all, alternate between the double object construction or taking a PP argument, as shown below:

- (8) a. John gave the book to Bill. [PP-dative]
 b. John gave Bill the book. [Double Object Construction]
- (9) a. John donated a painting to the museum. [PP-dative]
 b. *John donated the museum a painting. [Double Object Construction]

Even though *give* and *donate* are similar in meaning, the former can occur in either construction, while the latter only takes a PP argument. Although adults are aware of this restriction, children are found to overgeneralize the dative alternation to verbs that are incompatible in this frame (e.g., *don't say me that*, 3;2, Ross, MacWhinney corpus). Since children use language productively, they generalize argument structure properties

to verbs that could potentially follow a rule. Thus, children who assume that *donate* can occur in a double object construction, just like *give*, would eventually have to retreat from this superset hypothesis to the adult grammar where not all verbs alternate.

Throughout the dissertation, I discuss when children generalize from the input, and when they do not. I argue that overgeneralizations are only found when there is sufficient positive evidence for a rule in the input. Children retreat from overgeneralization when there is insufficient evidence in the input, as determined by the Sufficiency Principle (Yang 2016). The Sufficiency Principle and the threshold of generalization is discussed further in Section 1.2.1.

1.2. Conceptual Framework

In this section, I describe two learning models that I implement and defend throughout the dissertation: the Sufficiency Principle (Yang 2016) and the Active Mapping Model of Language Acquisition. I adopt the Sufficiency Principle to determine when a rule can be generalized over a verb class from the linguistic input. In order to determine how children acquire verb classes, I propose a new model of language acquisition, the Active Mapping Model, which describes how conceptual and structural cues are used by the learner. These models are described in detail in the subsections below.

1.2.1. The Threshold of Generalization

Throughout the dissertation, I defend the idea that a rule for any given linguistic class is not always generalizable. Rules are only generalizable when there is sufficient positive evidence in the input. I adopt the Tolerance Principle and the Sufficiency Principle (Yang 2005, 2016) as a way of determining how much positive evidence in the input is sufficient to generalize a rule, or to say that a rule is productive in language. The aforementioned principles allow us to calculate the number of positive members required to generalize a

rule. The formulation of the Tolerance Principle and the Sufficiency Principle is provided in (10) and (11), respectively.

- (10) “The Tolerance Principle: If R is a productive rule applicable to N candidates, then the following relation holds between N and e , the number of exceptions that could but do not follow R :

$$e < N \text{ where } N := N/\ln N \text{ (Yang 2016:10)}$$

- (11) “The Principle of Sufficiency: Let R be a generalization over N items, of which M items are attested to follow R . R can be extended to all N items iff:

$$N - M < N \text{ where } N := N/\ln N \text{ (Yang 2016:140)}$$

The formulas above essentially state that in order for a rule to be generalizable for a class of N members, there can be no more than $N/\ln N$ exceptions. If a rule has more exceptions than $N/\ln N$, then each member of the class will have to be lexically learned. The intuition behind the Sufficiency Principle can be captured by the following example: imagine that we are on an expedition on the Galápagos islands, where we spot 10 new creatures of a certain kind. We notice that 8 out of 10 of these creatures fly; therefore, we may conclude that all creatures of this type fly. In this case, we may extend our conclusion for the 8 flying creatures that we have observed to the remaining two, even if we only saw them on the ground. In contrast, we would not reach the same conclusion if we noticed that only 2 out of the 10 creatures fly (example adapted from Charles Yang, p.c.).² In the dissertation, I defend the idea that a rule is generalizable iff there are enough positive members following a rule.

Both the Tolerance Principle and the Sufficiency Principle allow us to determine the number of positive members needed within a class for a rule to be generalizable. However, there is a crucial difference in how the two formulas are used. The difference between the

¹The reader is referred to Yang (2016) for details regarding the derivation of the Tolerance Principle and the Sufficiency Principle.

²See Yang (2016) for other linguistic applications of the Tolerance Principle.

two can be expressed in terms of positive or negative exceptions observed in the input. For instance, in a class with 10 members, if the learner has observed a property p for 8 of those members, and property q for the other 2 members, then the learner can conclude that the rule p is generalizable to an 11th member of the class ($2 < N/\ln N$). The Tolerance Principle, in this case, tells us that we would expect the learner to extend p for a new member of the class that the learner has not yet encountered. Moreover, the property q for the remaining two members will need to be learned as lexical exceptions. In contrast, the Sufficiency Principle applies when for a class of 10 members, the learner observes p for 8 of those members, but has no evidence for whether the other 2 members show p . The learner is unable to determine whether they have not observed p for those 2 members because p does not apply to them, or whether they have simply not had the opportunity to witness p . In this case, the Sufficiency Principle determines that the learner generalizes p to the remaining 2 members of the class. In other words, the Sufficiency Principle determines whether enough positive evidence has been observed for a rule to be generalizable to all members of a class.

The difference between the Tolerance Principle and the Sufficiency Principle is important in addressing the errors of overgeneralization in child speech. These errors occur when the learner does not observe a property for certain members of a class, but extends this property to those members based on the observation that other members of the class show this property. In the dissertation, I investigate the properties of verbs, and verb classes, and ask whether a rule for a given class of verbs is generalizable to the other verbs within the same class. In the cases described in the dissertation, the Sufficiency Principle is invoked, as the learner does not encounter positive exceptions to the rule. Here, the learner must make the decision whether the evidence they have encountered in favor of the rule is enough to make a generalization that extends to the other members of the class.

To consider a hypothetical example, let us assume a class with 10 members where 4

members show a certain property. If we follow the Sufficiency Principle, 4 out of 10 is insufficient evidence to assume that the rule is generalizable to the other 6 members in the class. On the other hand, if there is positive evidence for 9 out of 10 members for this property, there would be sufficient evidence for the learner to generalize within the class. Note that in this case, they have not encountered any positive evidence that the one remaining member does not follow the rule. This is not always the case. At times, the learner may encounter positive evidence indicating that some members do not follow the rule under consideration, as in the case of past tense formation that consists of some irregular verbs.³ The rule can still be generalizable in these cases. Let us assume for the class of 10 members that the learner has positive evidence that 9 follow a rule and positive evidence that 1 does not. In this case, the number of exceptions is below the threshold of the Tolerance Principle, and the rule is productive. Therefore, if the learner encounters a new 11th member of this class, they will assume that this member follows the rule in the absence of positive evidence. I follow this distinction between the Tolerance and Sufficiency Principle throughout the dissertation.

Finally, in order to understand the problem of overgeneralization in the cases considered here, verbs must be viewed as part of a larger class from which the overgeneralization may or may not arise. We are only faced with the task of retreating from an overgeneralization when a property within a certain class is generalizable to other members of the class. I show that this is the case for causatives and English passives, but not for control and raising verbs. When there is insufficient evidence for generalization, the property must be learned individually for each class member. I show that in the cases of overgeneralization, a rule can be generalizable early on in the acquisition process when the learner's vocabulary size is small, and then fail to be generalizable in later stages. This accounts for the retreat from overgeneralization that results in the grammar we find for

³It should be noted here that when learning the past tense in accordance with the Tolerance Principle, children also extend the general rule to irregular verbs that are initially produced correctly before the rule is learned. The child must then relearn the irregular rules once again. Since this phenomenon is not observed in any of the cases described in the dissertation, I do not address it any further.

adults. In this way, the problem of overgeneralization is resolved. Moreover, as rule-learning is a primitive part of language acquisition, if a rule is unproductive for a broad class of verbs, I argue that learners seek subclasses where there is a generalizable rule to be found.

An example of how the Sufficiency Principle can be used can be seen in the acquisition of the dative alternation (Yang 2016). Yang 2016 discusses the productivity of child errors with dative double object constructions, resulting in overgeneralization. Yang 2016 argues that children generalize the dative alternation in early stages of acquisition when a high number of verbs in their vocabulary occur in the input both with a PP indirect object and in a double object construction. Here, it is also argued that in later stages of acquisition, the number of verbs that undergo the alternation, as the learner’s vocabulary size increases, fall below the threshold of the Sufficiency Principle. Therefore, children retreat from their overgeneralization due to insufficient positive evidence. A summary of the dative alternation findings from Yang 2016 is provided in Table 1.1.

Table 1.1: Summary of the number of verbs that undergo the dative alternation in the adult grammar. The dative alternation is productive for the more frequent verbs, but it is unproductive when the more infrequent verbs are taken into consideration (Yang 2016:167)

Top	Yes	No	N	Productive?
10	9	1	4	YES
20	17	3	7	YES
30	26	4	9	YES
40	30	10	11	YES
50	34	16	13	NO
60	39	21	15	NO
70	43	27	16	NO
80	46	34	18	NO
92	50	42	20	NO

Table 1.1 shows that the dative alternation is frequent for the 40 most frequent verbs in the input. However, as the more infrequent verbs are included, the dative alternation ceases to be productive. The frequency of the verbs in Table 1.1 reflect the stages in the

learner's vocabulary acquisition. As shown in 1.1, when the learner's vocabulary size is small, they have likely only learned the more frequent verbs which do largely show the dative alternation. As their vocabulary size increases and they learn the verbs that are less frequent, the dative alternation drops below the threshold of generalizability, and the learner is conservative in assuming which verbs can occur both with a PP-dative and in a double object construction. Thus, using the Sufficiency Principle, we can determine when the rule is generalizable in the input, and whether the generalizability of the rule changes with the learner's vocabulary size.

1.2.2. Active Mapping Model of Language Acquisition

In the previous section, we saw that the Tolerance Principle and the Sufficiency Principle can be used to determine whether the learner generalizes a rule from the input. Under this approach, the learner makes a generalization over a verb class when acquiring verb argument structure. In this section, I describe a model of how the learner comes to acquire the verb classes assumed in order to apply the learning model previously described in Section 1.2.1.

To account for the language acquisition facts described in this dissertation, I propose a new model of language learning: the Active Mapping Model of Language Acquisition. The Active Mapping Model of language acquisition distinguishes itself from models that propose an unlearned mapping between the semantics and the syntactic structure. Instead, I propose that any mapping between the syntax and the semantics is in fact acquired through the learner's linguistic experience. Throughout this dissertation, and especially in Chapters 3 and 4, I show how the Active Mapping Model can be implemented to account for the developmental trajectory of the learner in acquiring verb argument structure.

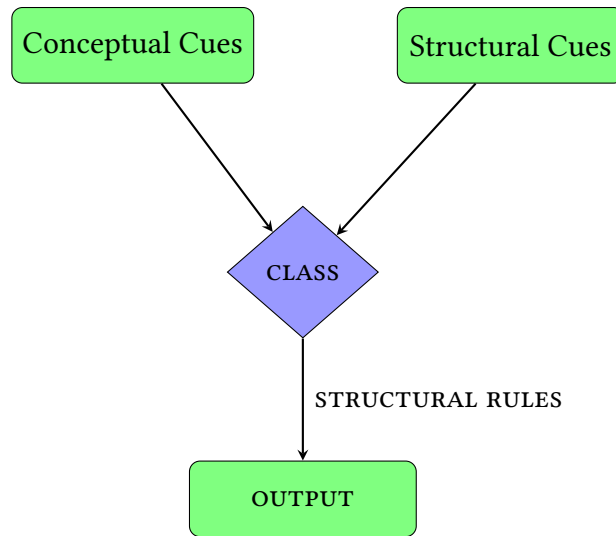


Figure 1.2: The Active Mapping Model of Language Acquisition does not have a built in link between the conceptual and structural cues; they are independent of each other. In this model, both conceptual and structural cues play a role in acquiring verb classes. Any rules that can be identified in the linguistic input apply at the class level.

The Active Mapping Model of Language Acquisition uses the conceptual and structural cues that the learner encounters in the linguistic input to form verb classes. The conceptual cues include both those cues that are observable from the learner’s non-linguistic environment and the word meaning, which is linguistic information that feed a class. The structural cues consist of the syntactic information in the linguistic input. The way in which these cues are used can be illustrated as follows. Let us assume that the learner hears a few simple transitive sentence in the input. Some examples are given below.

- (12) a. John saw the dog.
 b. The girl kicked the ball.
 c. The medicine helped Bill.

If this was the learner’s input, then the following structural observations could be made by the learner as part of the set of structural observations that the learner may make. Phrases such as *John*, *the girl*, and *the medicine* occur in initial position, phrases such as *the dog*, *the ball*, and *Bill* occur in final position, etc. (12b) illustrates a typical case of a transitive

sentence where the learner might observe that the sentence shows intentionality, and *the ball* is being acted upon by *the girl*. These would consist of some of the conceptual cues in (12b). In (12a), the learner might also observe the conceptual information that, for instance, there is no intentionality and that nothing is happening to *the dog*. For (12c), the conceptual cues could consist of no intentionality, and *Bill* being affected by *the medicine*. In this manner, the learner uses conceptual and structural cues available to them. The fact that *the dog* occurs in final position does not necessarily entail that it is tightly associated with a particular thematic role.

In this model, and throughout the dissertation, I use the term “conceptual cues” to refer to both situational cues that are obtained from the learner’s environment, and the semantics that may be obtained from the linguistic environment. Crucially, I do not dismiss the importance of linguistic cues in word learning. Rather, in the cases discussed in the dissertation, I emphasize the importance of situational cues separately from linguistic cues. For instance, one type of situational conceptual cue is that of intentionality. A long line of research has emphasized infants’ awareness of the goal-oriented nature of actions (Carey and Spelke 1994; Csibra, Bíró, Koós, and Gergely 2003; Wagner and Carey 2005; Woodward 1998). As previous work has shown that children as young as 4-6 months notice goal-directed actions, intentionality is adopted throughout the dissertation as a cue that is available to the learner early on.

I use the term “structural cues” to refer to the syntactic structure. Conceptual and structural cues both apply when forming a class; however, there is no innate mapping between the conceptual and structural cues. Conceptual cues only feed semantic information, and structural cues only feed structural information. Any mapping between the two may happen over time in the language acquisition process depending on the learner’s linguistic experience. The two types of cues are independent of one another. In order to form verb argument structure rules in their language, the learner only uses structural information from the input. These rules necessarily apply at the level of a verb

class. The distinction between the cues used to form a class must necessarily be distinct from the structural rules that apply at the class level.

The model proposed here is a response to the findings in the language acquisition and the theoretical linguistics literature that the correlation observed in regards to the semantics and the syntactic structure is not as tight a correspondence that an innate bootstrapping or linking rules model (e.g., Pinker 1989) would predict. For instance, Bowerman 1988 finds that children make causative errors with verbs whose semantics they are aware of early on. Children make overgeneralization errors with intransitive verbs that could potentially occur as a transitive causative, but do not in the adult grammar. Under an innate linking rules model, any exception of this kind cannot readily be accounted for. Throughout the dissertation, I discuss instances where the correspondence between the syntax and the semantics breaks down, and as a result, it cannot result from a learning model where the linking rules between the two are innate. By severing these two aspects of the grammar in the learning model, we can account for the exceptions to the correlations. Under this model, the learner arrives at some syntax-semantic mapping in their language by learning from the input. The mistakes made by the learner in the language acquisition process are also predicted under this model, as the learner is left to identify the mappings and the rules in their language without any guiding internal factor.

One question that this model raises is what are the kinds of conceptual and structural cues that are used by the learner. In this dissertation, I only begin to address this question by giving some indication of what these may be in the context of the cases examined in Chapters 2 – 4. In addressing this question, I show that structural cues are reliable cues, and are, therefore, exploited by the learner. Some conceptual cues, such as intentionality (Woodward 1998), may be available to the learner simply because we are attuned to them early on. I elaborate on these cues in more detail for the cases of the acquisition of causatives and passives in Chapters 3 and 4, and also discuss them throughout the dissertation.

The rest of this dissertation is structured as follows: Chapter 2 describes the acquisition of control and raising verbs, and argues that children learn these two constructions on the basis of non-referential subjects. Contrary to previous claims (Becker 2006), I show in this chapter that the problem of overgeneralization in learning raising and control verbs does not arise. Chapter 3 describes the acquisition of causatives, and in doing so, discusses how children acquire the class of unaccusative and unergative verbs in English. I demonstrate how the developmental trajectory seen in learning the causative alternation motivates the Active Mapping Model of Language Acquisition. Furthermore, I show that the overgeneralization of the causative rule arises and ceases in accordance with the threshold of generalization determined by the Sufficiency Principle (Yang 2016). Also in Chapter 3, I experimentally test the threshold of generalization proposed by the Sufficiency Principle. In doing so, I test the predictions of the Sufficiency Principle against entrenchment theories of verb argument structure learning (e.g., Ambridge et al. 2008). Chapter 4 discusses the acquisition of passives using the models described in Section 1.2. In particular, I address the asymmetry between actional and non-actional verbs in the acquisition of English passives (Maratsos, Fox, Becker, and Chalkley 1985). I demonstrate how this asymmetry is accounted for under the Active Mapping Model of Language Acquisition, as described in Section 1.2.2.

Chapter 2

The Acquisition of Control and Raising Verbs

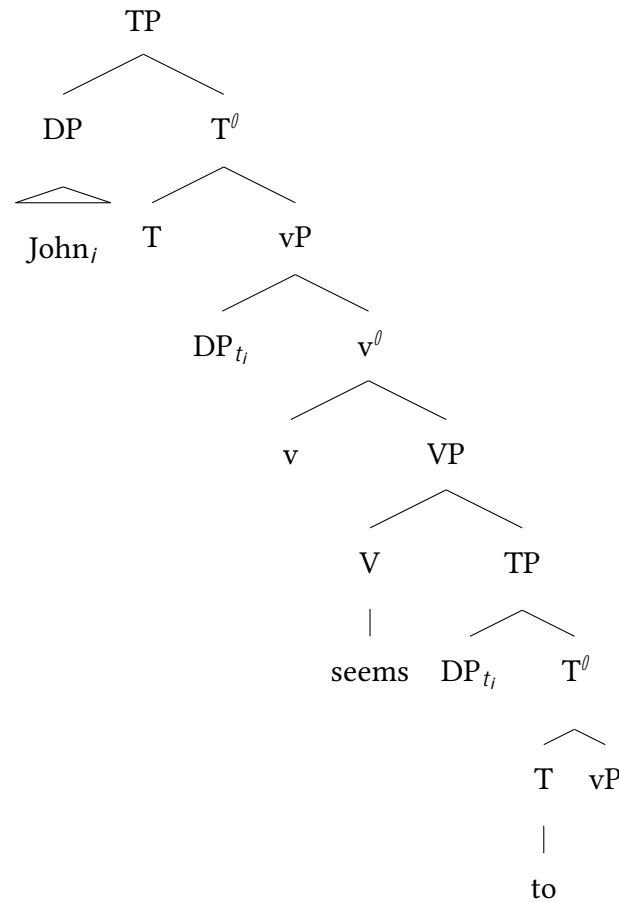
2.1. Introduction

Control and raising verbs are an integral part of linguistic theory. The question of how these verbs may be acquired is also an interesting one, as, on the surface, the two constructions can appear to be identical to the naïve learner.

- (13) a. $John_i$ wants [PRO_i to like syntax]. [Control]
b. $John_i$ seems [t_i to like syntax]. [Raising]
c. $John_i$ began [PRO_i to like syntax]. [Control]
d. $John_i$ began [t_i to like syntax]. [Raising]

However, in terms of their structure, the two constructions differ in crucial ways. In (13a), *want* is a control predicate, which means that the DP *John* does not raise from within the infinitival clause; we get a PRO subject there instead. A control structure is illustrated in (14).

(15) *Raising*



Given these three kinds of predicates, pure control verbs like *want*, pure raising verbs like *seem*, and ambiguous ones like *begin*, a learnability problem has been foregrounded in the literature (Becker 2006). Since some verbs, such as *begin*, are both raising and control, what would prevent the learner from concluding that all control verbs are also raising, and vice versa? Although the two verb classes are often distinguishable by the use of expletives, Becker 2006 argues that even though they are a useful cue in identifying raising verbs, they do not help to differentiate between pure raising verbs or ambiguous verbs. Thus, if the learner decides that some raising verbs are also control, what prevents them from concluding that all raising verbs are, in fact, verbs that can ambiguously take either a raising or a control structure? Once the learner has concluded that all verbs are ambiguous verbs, there is no positive evidence to guide the learner to the adult grammar, where there are verbs that can be purely raising or have either a raising or a control structure.

The question then is, how does the learner retreat from the superset grammar where they assume that verbs can take both structures, to a more conservative grammar where they can only take one? The retreat from overgeneralization is especially curious, given that direct negative evidence or feedback is unavailable or ineffective. This is a classic subset problem that has also been referred to as the problem of overgeneralization in the literature, which was described in Chapter 1. If children are indeed overgeneralizing from the hypothesis that all predicates could be both control and raising, the problem of retreating from this overgeneralization then arises, as the adult grammar also has predicates that can be only control or raising.

To address this potential problem, Becker 2006, 2014 proposes a learning model where retreat from overgeneralization occurs from the varying rates of inanimate subjects between the control and raising verb classes. Becker 2006 capitalizes on the fact that raising verbs allow both inanimate and animate subjects more frequently than control verbs, and uses this property to distinguish between control and raising predicates, as well as to distinguish ambiguous verbs from pure raising verbs. Control verbs are primarily predicted to occur with animate subjects, and therefore, verbs that are ambiguous between control and raising are predicted to occur with less inanimate subjects than pure raising verbs. The three kinds of predicates are, thus, argued to be learned through the different rates of inanimate subjects.

In earlier work, Becker 2006 claimed that a raising structure is the default structure for verbs that take infinitival complements, as opposed to a control structure with an embedded PRO subject in the complement clause. Becker reaches this conclusion on the basis of a series of experiments that tested children's comprehension of raising and control predicates, which we elaborate on in the next section. Due to the results from these experiments where control verbs appear to be initially analyzed as raising, Becker bases her model of the acquisition of raising and control verbs on the notion that the learner must retreat from an overgeneralization of a raising structure to all verbs that take an infiniti-

val complement. Hence, children then learn that a predicate is not raising if they show a lower rate of inanimate subjects than verbs that are classified as pure raising, which is typical for control verbs. This type of learning is a form of indirect negative evidence because children are said to generalize from the absence of inanimate subjects, as opposed to a positive occurrence of a linguistic cue.

In this chapter, I argue instead that control is acquired prior to raising, and that the problem of overgeneralization does not arise if the adopted theory of generalization is reasonably constrained. I also argue here that a theory of non-referential subject-based learning fares better than an animacy-based approach.

In the analysis I present here, only subject control and raising to subject verbs are considered. This decision was made for two main reasons. First, previous analyses of the acquisition of raising and control verbs (e.g., Becker 2006, 2014; Hirsch and Wexler 2006; Mateu 2016; Orfitelli 2012) primarily focused on subject control and raising to subject predicates. More is known in the literature about subject control and raising to subject constructions, and the analysis presented in this paper also builds on previous work on these constructions. The second reason for this decision is that there is evidence suggesting an asymmetry between the acquisition of object control and raising to object verbs versus the acquisition of subject control and raising to subject verbs (Chomsky 1969; Sherman and Lust 1993). Notably, C. Chomsky (1969) discusses the Minimal Distance Principle where children treat subject control verbs with multiple antecedents (e.g., *promise*) as object control. To avoid any confounds and to be able to draw parallels with previous work, only the latter kinds of verbs are included in the analysis.

The rest of this chapter is structured as follows: first, I review previous work on the acquisition of raising and control. Next, I discuss the plausibility of the aforementioned learnability puzzle and Becker's solution to the problem in Section 2.2, which is based on the rate of inanimate subjects found with the three kinds of predicates. In Section 2.3, I show that the difference in the proportion of inanimate subjects between the three verb

types is not salient enough for the learner to accurately acquire the distinct properties of these predicate classes. Instead, in Section 2.4, I propose that children learn the three classes of predicates from positive evidence in the linguistic input. In this case, it is the use of non-referential expletive subjects¹ that provides the learner with the cue needed to acquire raising and control verbs.

2.2. A Supposed Learnability Puzzle

Becker 2006 considers, and eventually rejects, the possibility that children initially expect a control structure when a verb occurs with an infinitival complement. If children initially assume that all verbs that take infinitival complements are control verbs, non-referential expletive subjects are one way to identify the verbs that are raising because they indicate that no theta-role has been assigned to the external argument introduced by a verb. Control verbs, unlike raising verbs, are incompatible with expletive subjects. However, Becker 2006 argues against this learning mechanism on the basis of verbs like *begin* that fall into both the control and raising categories. If the learner analyzes *begin* as a pure raising verb using expletive subjects, as in *it began to rain*, then the learner would have no positive evidence to also identify it as a control verb later on. If the learner assumes both a raising and a control structure for *begin*, then there is nothing in the input that prevents the learner from assuming that pure raising verbs like *tend* or *seem* are also optionally control verbs like *begin*. In other words, what prevents the learner from overgeneralizing the property that verbs can occur in both control and raising structures? The possibility of running

¹Throughout the analysis, I use the term “expletive subject” and “non-referential subject” interchangeably or in conjunction. The term “expletive” is used to also include the constructions in null subject languages in which the grammatical subject position is null. I am agnostic as to whether that position is occupied by an unpronounced expletive, or is simply absent. The account proposed in this chapter is also applicable to null subject languages as children learning these languages are attuned to the referentiality of the null content. Evidence for the knowledge of the content of null subjects at an early age is evident, for instance, from work on Mandarin, a topic drop language. Wang, Lillo-Martin, Best, and Levitt (1992) show that children between the ages of 2 and 4 never use an overt non-expletive lexical item in weather constructions, even when prompted to do so. The fact that non-referential subjects in their language are unpronounced, thus, appears to be known by children learning pro-drop and topic-drop languages.

into the problem of overgeneralization is one reason why Becker 2006 assumes that verbs that take infinitival complements are initially analyzed as having a raising structure.

Becker 2006, 2014 contend that any analysis of the acquisition of raising and control predicates must address the problem of learning pure raising, pure control, and ambiguous predicates without overgeneralizing that all predicates are ambiguous between raising and control. A visualization of the problem can be seen in Figure 2.1 below:

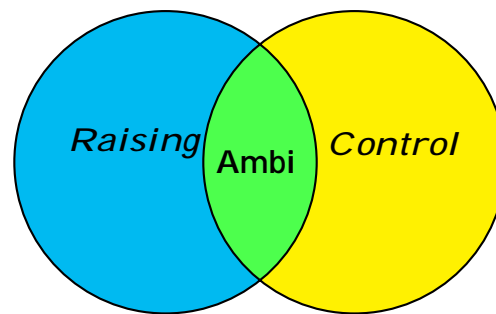


Figure 2.1: Relationship between control, raising, and ambiguous verbs.

As illustrated in Figure 2.1, the problem any analysis of the acquisition of control and raising verbs faces is that of singling out pure control and raising verbs, while still allowing for a class of verbs that fall into both categories. In doing so, we must additionally take care to prevent a situation in which all potential control and raising predicates are generalized as being ambiguous predicates.

The problem of overgeneralization is one reason that leads Becker 2006 to hypothesize that children initially assume all verbs that take an infinitival complement are raising verbs. Another reason for assuming a raising first analysis is that Becker 2006 conducted a series of experiments using truth value judgment tasks which showed at first glance that children appear to analyze *want* as a raising verb. The studies find that children tend to treat sentences like *the ower wants to be pink* (Becker 2006:445) as a raising construction because they rate this sentence as pragmatically acceptable. Adult controls, in contrast, tend to judge such sentences as infelicitous or “silly”. These studies interpret children’s acceptance of these sentences to be the result of them assigning a raising structure to *want*.

The results from this experiment, and the reasoning on the problem of overgeneralization, together lead Becker 2006 to conclude that children first assume a raising structure when they encounter a verb that takes an infinitival complement. In later work, Becker 2014 refines this claim to argue that children only assume a raising structure when they hear an inanimate subject with control and raising verbs. However, this refinement still postulates that in a sentence like *the ower wants to be pink*, children extend a raising structure to *want*, which is a control verb.

The results of Becker 2006's experiments suggest findings contra work that show the acquisition of control at an early age (Cairns, McDaniel, Hsu, and Rapp 1994; Goodluck 1981; Goodluck, Terzi, and Díaz 2001; Hsu, Cairns, and Fiengo 1985; Maratsos 1974; McDaniel, Cairns, and Hsu 1990/91; Sherman and Lust 1993). In previous studies, Sherman and Lust 1993, for instance, finds that in sentences without multiple linguistic antecedents, children as young as 3 years old pass comprehension tests.² Additionally, in crosslinguistic work, Goodluck, Terzi, and Chocano Diaz 2001 illustrate that 4-5 year old children show the ability to comprehend control sentences in Spanish and Greek. However, as Hirsch and Wexler 2007 note, Becker's results could also be interpreted as younger children treating and accepting *the ower* to be animate for the purposes of the task. Although the results are also consistent with children interpreting *want* as a raising verb, the possibility of another explanation indicate that the results cannot be taken to be inarguable proof for assuming a raising structure default.

As mentioned earlier, in subsequent work, Becker 2009, 2014 depart from the explicit support of a raising-first hypothesis, but maintain the argument that the learner assumes a raising structure for predicates that occur with an inanimate subject. This argument

²The difference between single and multiple antecedents is crucial in determining children's ability to comprehend control sentences. For instance, in a sentence like *John promised Bill to go to the store*, the null subject PRO is coindexed with *John*, and not *Bill*. Children are known to misinterpret these sentences and assume incorrectly that PRO is coindexed with *Bill* (Chomsky 1969). Children in these cases do not pass comprehension tests as they pick out the wrong antecedent for PRO. However, children pick out the correct referent when there is one linguistic antecedent present. These facts indicate that children can generally comprehend control structures, but that the presence of multiple antecedents introduces further complications.

means that the results of the experiment can be seen as children accepting *the over wants to be pink* as raising because children have not yet learned that *want* only takes animate subjects; i.e., they have not learned that *want* is a control predicate. However, considering the evidence that children have mastered control constructions well before raising constructions (e.g., Goodluck, Terzi, and Díaz 2001; Sherman and Lust 1993) and Becker 2014's retraction of the raising-first hypothesis, I assume in this chapter that children have a default control structure for verbs that take an infinitival complement. I assume a control default even for cases where a control verb like *want* occurs with an inanimate subject. The raising structure of verbs, in contrast, must be learned.

The rest of this chapter argues that the problem of overgeneralization for ambiguous predicates does not arise, and as a result, we need not assume that the learner initially assigns a raising structure to the verb, even when it occurs with an inanimate subject. I first provide a quantitative evaluation of Becker 2014's indirect negative evidence approach in the following section, Section 2.3. In the next section, I show that the rates of inanimate subjects with the three kinds of predicates are not different enough for the learner to disambiguate between them accurately.

2.3. An Indirect Negative Evidence Approach

In this section, I show that the rate of inanimate subjects in the linguistic input does not differ enough between the three kinds of predicates to allow for learning from indirect negative evidence. I show that the data in the input is not adequate for indirect negative evidence to serve as the primary learning mechanism. In the following subsections, I review the plausibility of children learning from the tendency of raising predicates to select for more inanimate subjects than control predicates. I first present the rate of inanimate subjects with the three classes of predicates in question as reported by Becker 2014, and review the plausibility of learning from indirect negative evidence given the rate of

inanimate subjects found by Becker 2014. Next, in Section 2.3.2, we perform a backwards calculation to determine the amount of input data required by the learner to learn the properties of a given predicate.

2.3.1. Rate of Inanimate Subjects with Control and Raising Verbs

Becker 2014 determines the rate of inanimate subjects with raising, control, and ambiguous predicates by examining the mother’s speech for Adam, Sarah, and Eve in the Brown (1973) corpus in CHILDES (MacWhinney 2000). Tables 2.1-2.3 (from Becker 2014, Table 6.1) list the number of animate and inanimate subjects that occur with each type of predicate. We see in these tables that all three types of predicates occur overwhelmingly with animate subjects, and my own examination of the data also reveals the same.

Let us first take a look at Becker 2014’s data for raising predicates in child-directed speech. In Table 2.1, we see that the overall rate of the use of inanimate subjects is 5.2%. We also see that two out of the three verbs, *used to* and *going to*, are used with a low percentage of inanimate subjects, especially when compared to the overall rate of inanimate subjects of ambiguous verbs as shown in Table 2.3.

Table 2.1: Inanimate subjects with raising verbs (from Becker 2014, Table 6.1)

Raising	Animate	Inanimate	% Inanimate subjects
<i>seem</i>	4	6	60%
<i>used (to)</i>	45	5	10%
<i>going (to)</i>	1197	58	4.6%
total	1246	69	5.2%

Under an animacy-based approach, we might expect that inanimate subjects should occur with raising predicates and ambiguous predicates significantly more frequently than with control predicates. However Table 2.1 shows that many raising verbs occur primarily with animate subjects.

Additionally, it can be seen that a large proportion of the inanimate subjects that are found for the class of raising verbs occur with *seem*. *Seem* occurs with inanimate subjects at a rate of 60% in the Brown corpus. The total number of inanimate subjects with *seem* is also likely higher than in child-directed speech more generally. This is confirmed when we look at all the input data available for North American English learning children in CHILDES. Overall, we find that *seem* occurs with inanimate subjects at a lower rate of around 30%. This lower rate of inanimate subjects with *seem* in the overall CHILDES data would then likely bring down the rate of inanimate subjects with the class of raising verbs, making it closer to that of the control class shown in Table 2.2.

Table 2.2 shows that control predicates also primarily occur with animate subjects in the input. Both Table 2.1 and Table 2.2 illustrate that control verbs and raising verbs both occur primarily with animate subjects. The overwhelming presence of animate subjects still holds for both raising and control verbs, although there is still some observable difference in the distribution of the rate of inanimate subjects between the two kinds of predicates.

Table 2.2: Inanimate subjects with control verbs (from Becker 2014, Table 6.1)

Control	Animate	Inanimate	% Inanimate subjects
<i>want</i>	405	2	0.5%
<i>like</i>	210	0	0%
<i>try</i>	86	0	0%
<i>love</i>	10	0	0%
<i>hate</i>	1	0	0%
total	712	2	0.3%

A Fisher's Exact test on the proportion of inanimate subjects to animate subjects to reveal a significant difference between the classes of pure raising and pure control verbs ($p < 0.01$). However, the question addressed in the rest of this chapter is not only whether there

is a significant difference between the proportion of inanimate subjects between raising and control verbs, but also whether the difference is usable by children given the learning model. As discussed later on in this chapter, we see that the number of inanimate subjects is not always significantly different, and moreover, given the learning model proposed by Becker 2014, there is not enough evidence in the input for the difference to be usable by the learner.

Finally, given Becker 2014’s analysis, the rate of inanimate subjects for ambiguous predicates may be expected to fall in between that of control and raising predicates, but that prediction is not borne out. Instead, we find that the class of ambiguous predicates is not readily distinguishable from the pure control and pure raising classes.

Table 2.3: Inanimate subjects with ambiguous verbs (from Becker 2014, Table 6.1)

Ambiguous	Animate	Inanimate	% Inanimate subjects
<i>start</i>	4	0	0 %
<i>begin</i>	1	0	0%
<i>need</i>	38	4	9.5%
total	43	4	8.5%

Table 2.3³ shows that the ambiguous predicates *start* and *begin* are vanishingly rare in the data, and they do not occur with inanimate subjects. Moreover, the overall use of inanimate subjects is only 8.5%. It seems numerically unclear whether there is a large enough difference between the classes for the learner to tell them apart. Statistically, however, the rate of inanimate subjects with pure raising verbs is not significantly different from that of ambiguous verbs. Fisher’s Exact test on the overall use of inanimate subjects with pure raising verbs and ambiguous verbs reveals that the two overall proportions when compared to each other are not significantly different ($p = 0.3$). This means that the numbers

³Although *begin* occurs infrequently in the Brown corpus, I show later on that its frequency is in fact much higher in CHILDES overall.

noted for the ambiguous predicates are not distinct from the numbers for the pure raising predicates.

Furthermore, Tables 2.1-2.3 indicate that the rate of inanimate subjects varies greatly between predicates. A number of predicates in the control class and the ambiguous predicates class, for instance, do not occur with inanimate subjects at all. In sum, there is a tendency for raising predicates to occur with inanimate subjects, but most verbs do not occur very frequently in the data, and when they do, they occur mostly with animate subjects.

What we have seen thus far is that the animacy approach claims that children note the predicate's level of selectivity for animate subjects, and conclude that the higher a predicate's selectivity for taking animate subjects, the greater the likelihood of it having a control structure. This analysis is based on the premise that children form a generalization when they do not see inanimate subjects sufficiently frequently with certain predicates. This proposal falls under the umbrella of indirect negative evidence due to the fact that the learner is said to generalize based on the lack of evidence. As a result, an indirect negative evidence approach must specify the environments in which overgeneralizations are not made (cf. Pinker 1989, Yang 2015). Learning control and raising predicates under this kind of indirect negative evidence approach proves to be a non-trivial task, as the conditions in which overgeneralizations are made are not specified. Moreover, even if learning to retreat from indirect negative evidence approaches is possible in principle, the evidence in the input is too sparse for the cues to be useful to the learner. Using the case of the acquisition of raising and control verbs, I show in the following section that learning these predicates from indirect negative evidence is not straightforward. I then provide an account of a positive evidence based approach of the acquisition of raising and control verbs in later sections.

As seen earlier, the number of inanimate subjects with all three types of predicates is small. An indirect negative evidence approach must show that these statistical disparities

are prominent enough in the input to be applicable for use to the learner. To determine whether the statistical differences are large enough to be usable by the learner, I estimated the probabilities of being able to distinguish these classes of verbs using indirect negative evidence by using the numbers provided in Becker 2014. However, we note here that the learner does not know a priori the percentage of inanimate subjects with the three kinds of predicates; an actual learning model must estimate the probabilities and use it to draw statistical inferences at the same time. This kind of learning would amount to the best case scenario for an indirect negative evidence learner.

2.3.2. The Probability of Learning from Indirect Negative Evidence

For the purpose of the argument, let us assume that children are aware of the rates of inanimate subjects with control and raising verb classes. Then, the question is, can the child use such differences in the attested usage of inanimate subjects to draw reliable conclusions about which class a given predicate belongs to?

We can decide whether the learner is able to draw conclusions based on indirect negative evidence by performing some statistical calculations on the learner's input data. In order to estimate the learner's input, I use child-directed speech attained from combining all the corpora on North American English available on CHILDES as of May 2017. Throughout the rest of this chapter, this combined data serves the purposes of gauging the input available to the learner in early stages of acquisition.

The task of distinguishing between the three kinds of predicates using the rates of inanimate subjects is analogous to the task of trying to distinguish between two coins with different probabilities of showing heads. In the coin example, we are tasked with distinguishing between 2 coins by flipping them repeatedly and sufficiently many times so that their differential probabilities can be reliably detected in the sample. Assuming that all trials are independent, intuition should make clear that the probabilities should be very close to the expected value, so the child can indeed know which class the predicate

belongs to. However, the number of times a predicate has to be heard in order for that to happen is unclear, as the sample size must be large enough in order for the learner to reliably identify the class of a predicate.

Let us consider a hypothetical example with two coins A and B, whose probabilities of showing heads are 0.2 and 0.5 respectively, and our task is to guess which coin is which. This example is similar to the learner differentiating between two types of predicates. If we flip the coin once, regardless of whether the outcome is heads or tails, we will not be able to reliably guess the nature of the coin; i.e., one instance of a predicate is not enough to reliably say which class the predicate belongs to, the raising class or the control class. The more we flip the coin, the more likely it is that we will be able to confidently guess whether we have coin A or B. If we flip the coin very many times, the probability distributions will cluster around the true means of the two coins. The task to distinguish between the two coins, on the other hand, cannot be done effectively with only a few coin tosses. Under the animacy-based approach, the learner is faced with the same kind of problem, but with values that are closer together than in the examples, which are described below. The learner has to hear a predicate a number of times in the input before they can determine the type of the predicate.

Now returning to the case at hand, as the learner, we would like to determine, given the number of inanimate subjects used, whether the predicate is raising, control, or both over N number of trials. As in the coin example, generally speaking, the more times we see a predicate used either with an animate or inanimate subject, the more likely we are to be able to confidently tell the predicates apart. Before we tackle the cases in question, let us consider a toy example where the learner can reasonably be expected to learn from indirect negative evidence. Let us suppose that we are trying to determine whether a Predicate X is of type A with the rate of inanimate subjects at 0.2, or whether it is of type B with the rate of inanimate subjects at 0.5. Taking these rates of inanimate subjects into account, we can calculate the probability that a predicate is either of type A or B.

At each instance of encountering the predicate, we will either see it with an animate or inanimate subject. This process is binomial, and therefore, we can use the binomial test in order to determine which occurrences of a predicate will be informative to the learner. A binomial test is performed at each trial to check whether the probability of the predicate being of type A is significantly different from the probability of the predicate being type B; i.e., whether at any point it could be determined that predicate X is of one type or the other. The probabilities of the predicate are then summed up at each trial where the result was significant. The result is the overall probability of being able to reliably determine whether predicate X is of type A or B, given the number of times we have encountered it. For instance, if we encounter predicate X 10 times, we get the numbers shown in Table 2.4.

Table 2.4: $P(i, N, a)$ and $P(i, N, b)$ indicate the probability of a predicate being of type A with the rate of inanimate subjects at 0.2 or of type B with the rate of inanimate subjects at 0.5 at each trial. i indicates the number of inanimate subjects a predicate has occurred with. A binomial test indicated by $bt(i, N, a)$ and $bt(i, N, b)$, shows whether the number of times a predicate occurred with an inanimate subject was significant to tell whether a predicate is of type A or B. The rows where the binomial test showed significance are presented in bold. The sum of the probabilities when the binomial test showed $p < 0.05$ reveals the cumulative probabilities over the total number of trials.

i	$P(i, N, a)$	$bt(i, N, a)$	$P(i, N, b)$	$bt(i, N, b)$
0	0.107	0.228	0.001	0.002
1	0.268	0.698	0.01	0.021
2	0.302	1.0	0.044	0.109
3	0.201	0.43	0.117	0.344
4	0.088	0.121	0.205	0.754
5	0.026	0.033	0.246	1.0
6	0.006	0.006	0.205	0.754
7	0.001	0.001	0.117	0.344
8	0.0	0.0	0.044	0.109
9	0.0	0.0	0.01	0.021
10	0.0	0.0	0.001	0.002

In Table 2.4, we see varying probabilities of being able to determine whether a predicate is of type A or B. For instance, if we encounter this predicate of an unknown type with

four inanimate subjects, we get the probabilities shown in row 4. Row 4 shows that the probability of predicate X being of type A is 0.088, as indicated by the $P(i, N, a)$ column. The probability of X being of type B is 0.205. The probability of X being of type A is not low enough to be able to reliably say that it is not of type A. However, in row 5, the probability of the predicate being of type A is low enough. As illustrated by the binomial test, in this instance, the learner is reliably able to identify whether the predicate is of type A or B. These instances are cases where we can assume that the data are informative to the learner. The rows for which the learner can reliably detect that X is either a type A or type B predicate are presented in bold. If we add up the probabilities of the rows in boldface, we find that if a predicate is of type A, there is only a 41% chance that we can detect it. There is a 62% chance of detecting the predicate type if it is of type B. With only 10 instances of inanimate subjects, the learner cannot reliably say which class the predicate falls under. In this example, the predicate type can reliably be detected if we encounter the predicate with 37 inanimate subjects, as shown in Figure 2.2.⁴

With 37 instances needed to reliably determine the predicate type, in this toy example, children can reasonably be expected to learn from indirect negative evidence when the probabilities are 0.2 and 0.5. These rates are different enough from each other that the number of instances of inanimate subjects required is relatively low.

Figure 2.2 shows that the probability of determining predicate type increases as the number of trials increases. Note that there are two variables at play in the figure: one variable is the number of times the learner hears the predicate, i.e., the sample size. The other variable is the number of times the predicate is heard with an inanimate subject. For any given number of times a predicate has been heard, there is a range of inanimate subjects with which it could have occurred. The graph presented here averages over the

⁴I am not claiming here that the learner is actively performing a binomial test when encountering the data in the input. Rather, the calculations presented here serve the purpose of illustrating how much data the learner would need to reliably make a decision regarding the properties of a predicate using the rate of inanimate subjects. The binomial test is used because not all instances of hearing the predicate will be equally informative to the learner. These calculations are simply a model of how learning from indirect negative evidence might take place using statistical cues and inferences.

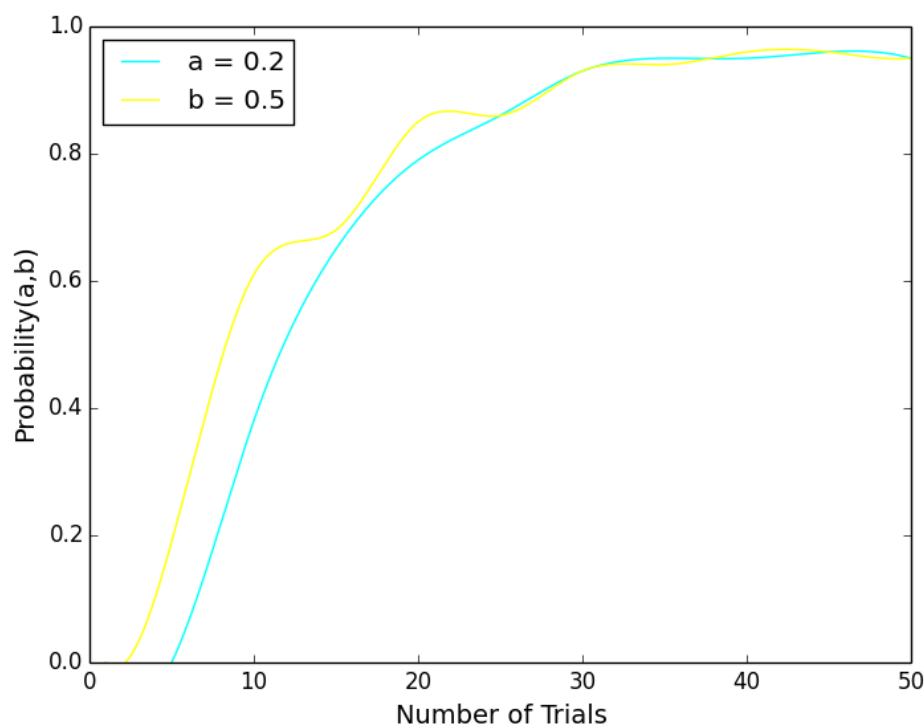


Figure 2.2: Number of occurrences of a predicate needed to determine whether the predicate is of type A or B.

number of times an inanimate subject occurs with a predicate. A range of inanimate subjects were included in these calculations because we cannot be certain of the number of inanimate subjects that the learner will encounter. The dips in the graphs occur when the probability of identifying predicate type differ due to the number of inanimate subjects heard with the predicate.

As shown earlier, the number of inanimate subjects and the number of trials required in the toy example are low. This is because the hypothetical probabilities of 0.2 and 0.5 are different enough from each other. However, in learning raising, control, and ambiguous predicates, the rates of inanimate subjects do not differ much between the classes. To illustrate this point, I discuss two cases: distinguishing the class of pure raising predicates from pure control predicates, and distinguishing pure raising predicates from ambiguous predicates. Here, we see that the statistical differences in the rate of inanimate subjects

are not different enough for the learner to reliably detect the predicate type.

Now that I have illustrated how the different rates of inanimate subjects can be used in a toy example, we can turn to the actual problem at hand, which requires the same process as described for the toy example. In order to estimate how many examples a learner would need in order to learn verb classes, we can assume a model similar to the one described above. Assuming $a = 0.003$ for the rate of inanimate subjects expected with control predicates and $b = 0.052$ for the rate of inanimate subjects expected with raising predicates (from Tables 2.1 and 2.2), a binomial test can be performed to find the total number of instances of any given predicate a child would have to hear in order to reliably distinguish the classes. Essentially, we are calculating the number of trials needed in order for the child to be certain whether a predicate is raising or control, or an ambiguous or pure raising predicate. Given that the rate of inanimate subjects for these classes are fairly close, a few instances alone are not enough to accurately determine which class the predicate belongs to. Similar to the examples presented earlier, the more instances of a predicate the child hears, the more reliable the distinction between the classes.

The result of performing these calculations on the control, raising, and ambiguous predicates shows that about 190 instances are needed to disambiguate between pure control and pure raising with a 99% confidence interval. With a 95% confidence interval, the child would need to hear about 90 instances of a predicate to make such a decision, as illustrated in Figure 2.3. In other words, the probability of confidently determining which predicate the learner has encountered approaches 1 when the learner has heard the predicate about 90 times.

In distinguishing between the pure raising and ambiguous predicates class, as the rate of inanimate subjects for pure raising would be $a = 0.052$, and the rate of inanimate subjects for ambiguous predicates would be $b = 0.085$ (Table 2.3). In this case, with a 99% confidence interval, the number of instances of a predicate the child must hear to differentiate between the classes is 1258. With a 95% confidence interval, at least 633

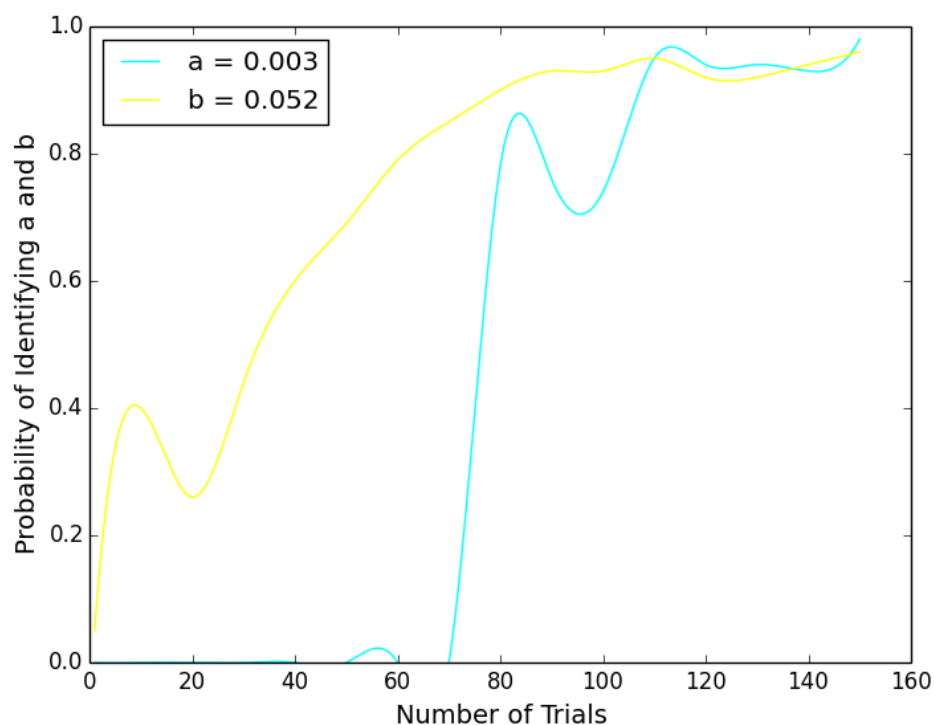


Figure 2.3: Number of occurrences needed to distinguish pure control predicates from pure raising predicates with a 95% confidence interval.

instances of each predicate has to be observed. The probability of being able to distinguish between pure raising and ambiguous predicates when the predicate has occurred N times with a 95% confidence interval is shown in Figure 2.4.

As Figure 2.4 illustrates, a predicate has to be heard several hundreds of times in the case of pure raising versus ambiguous predicates to be able to determine which class the predicate falls into. As we recall from Tables 2.1-2.3, many of the raising and control predicates do not occur that frequently in the data.

Using data from the combined corpora of English-learning children on CHILDES, we can show that even the 633 instances with a 95% confidence interval are indeed more data than the learner receives in the input. Since Becker 2014's analysis is based on the data in the Brown corpus alone, I use the combined input data sample, which is much larger, to illustrate precisely how much time it would take the learner to accumulate 1258

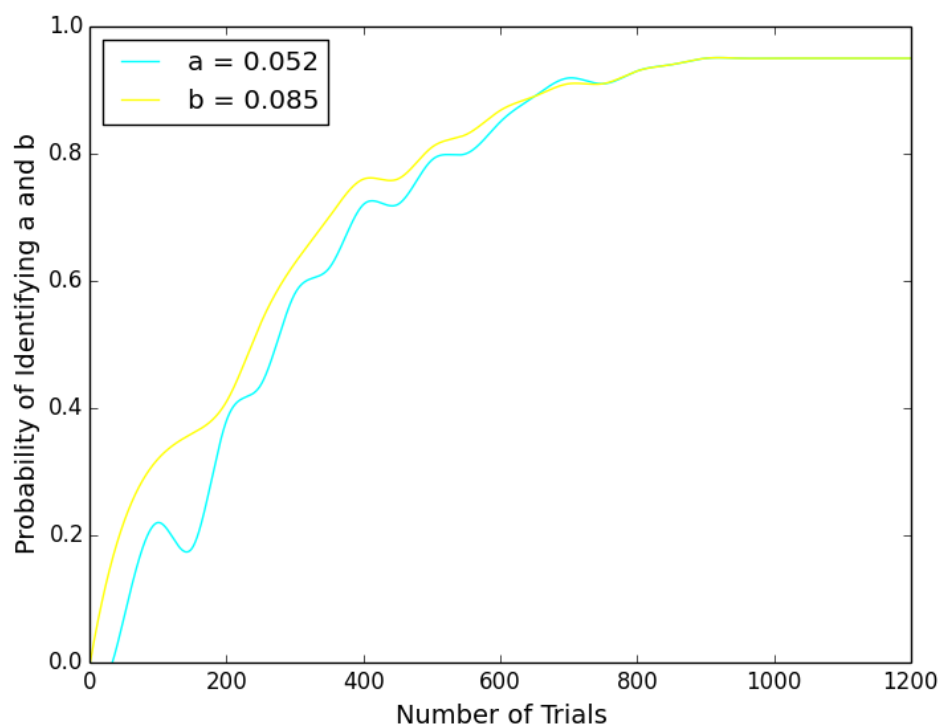


Figure 2.4: Number of occurrences needed to distinguish ambiguous predicates from pure raising predicates with a 95% confidence interval.

instances of the various predicates. The combined input data results in 6 million words of child directed English, which is roughly a year of speech, although we do expect linguistic input differences to vary between children (Hart and Risley 1995). In examining this combined input data, we find that there is a wide distribution of the frequency of ambiguous predicates. Some ambiguous predicates like, *have* and *need* are extremely frequent in the input, occurring approximately 47,257 and 8,714 times, respectively. There is enough evidence for the learner to learn these verbs are other extremely frequent ones under an indirect negative evidence model. However, 3 out of 11 of the ambiguous verbs examined occur less than 50 times in the input sample examined. These verbs, *manage*, *fail*, and *threaten*, only occur 38, 14, and 12 times. A positive evidence based approach has the merit that, broadly speaking, each and every datapoint is reliable (e.g., Marcus 1993). Indirect negative evidence based feedback is noisy as each datapoint is not unambiguous,

and therefore, the frequency of the verbs in the input matter. Since the two approaches make different predictions regarding verbs that are not extremely frequent in the input, I consider the two most infrequent verbs in this sample of 11 verbs, *fail*, which occurs about 14 times in the 6 million word data set, and *threaten*, which occurs 12 times in the data set. I examine *fail* and *threaten* to illustrate the predictions made by an indirect negative evidence approach for these verbs.

Examining the production data of North American English-learning children in the CHILDES database shows that there are instances of children using *fail* and *threaten* as verbs before the age of 8. Some examples from the child production data are provided below.

- (16) a. you won't fail (4;9, Karen, Hall corpus)
- b. I threatened them (6;7, Ross, MacWhinney corpus)
- c. and here's one that won't fail (7;7, Ross, MacWhinney corpus)
- d. but the aliens tried to threaten her (7;10, Gillam corpus)

As illustrated by the examples above, both *fail* and *threaten* are used by children with animate subjects. Moreover, *fail* also occurs with an inanimate subject, which is determined by examining the context in which the sentence was uttered. For *threaten*, we can see that the learner is aware of the control properties from the agentive and intentional interpretation of the sentences. Given this data, it seems likely that the learner is aware that *fail* and *threaten* can introduce an external argument and assign theta-roles to them. The uses of these verbs before age 8 suggests that the learner knows their control properties before adolescence at least. These uses would be unexpected if the learner needed to hear a large number of instances of these verbs used in order to determine that they can optionally be used as control verbs.

Although the production data suggests that children acquire *fail* and *threaten* fairly early on, an indirect negative evidence learner would not acquire these verbs until much

later. Assuming that children hear about 6 million words of speech per year on average (Hart and Risley 1995), which is about the size of the combined data examined, we would expect *fail* and *threaten* to be properly learned only around age 45 if the learner was indeed using indirect negative evidence. Even though a number of verbs such as *have* and *need* do occur thousands of times, there are many verbs in the input that do not. This frequency distribution follows from the fact that language follows Zipf's law. There are many verbs that do not occur often in speech, but the child is still able to learn the properties of a lexical item from few occurrences. If we assume that the learner needs 1258 instances to correctly identify the properties of a predicate, we would predict that many of these verbs do not get learned until much later in life, or not at all. In this analysis, the 11 ambiguous verbs were already selected based on their relatively high frequency; a number of ambiguous predicates that did not occur in CHILDES were excluded.

In Section 2.4, I propose an alternative solution to an animacy based approach: generalization from sufficient positive evidence. I adopt Yang 2016's Sufficiency Principle, outlined in Chapter 1, as a mechanism to determine when the learner generalizes a rule. The Principle of Sufficiency shows us that there is no condition under which ambiguous predicates are overgeneralized; therefore, the learnability puzzle never arises in the first place. Moreover, I argue that pure raising predicates are learned from non-referential subjects, which unambiguously differentiate themselves from control predicates. Ambiguous predicates are learned separately from pure raising predicates using the relative proportion of non-referential subjects they occur with in the input. Ambiguous predicates occur with fewer non-referential subjects than pure raising verbs.

2.4. Generalization from Positive Evidence

In this section, I demonstrate how the learner arrives at the distinctions between raising and control using positive evidence in the input. I show that positive evidence in this case

manifests itself in the form of non-referential expletive subjects. I aim to show here that non-referential subjects are a more reliable indicator of the difference between the three types of predicates than subject animacy. I also show that the supposed learnability puzzle resulting from the overgeneralization of ambiguous predicates does not arise. Therefore, the problem of overgeneralization never occurs, and the assumption that children analyze verbs that take infinitival complements as control structures initially is not theoretically problematic.

Sections 2.4.1 and 2.4.2 discuss in detail how the problem of overgeneralization does not arise in learning control and raising predicates. In these sections, I first argue for the use of non-referential subjects as indicators of the differences between control and raising predicates. Next, I provide the number of ambiguous, raising, and control predicates that occur in the input to show that the number of ambiguous or raising predicates are not enough to be generalizable by the learner. In Section 2.4.3, I describe how ambiguous predicates are learned distinctly from pure raising predicates through the relative rates of non-referential subjects.

2.4.1. Positive Evidence for Learning Control and Raising Verbs

This section describes the nature of non-referential subjects in the input, and how children use this data to learn raising and control predicates. I propose the following learning mechanism: potential raising or control predicates are first identified when the verb takes an infinitival complement. The distinction between the raising, control, and ambiguous verb classes are, thereafter, learned through the use of expletive subjects. If a predicate is used with an expletive, this cue is used by the learner to assume that the predicate can take a raising structure.

The presence of expletive subjects is an indicator of a crucial aspect of raising constructions: these verbs do not assign a theta-role to their subject. Moreover, when expletive subjects occur with raising predicates with a CP complement, the subject is low,

as in *It seems that John likes syntax*. Such examples serve as evidence to the learner that subjects of raising verbs originate low. All the known properties of raising predicates that differ from control are derivative of these facets of raising verbs. The analysis presented in this dissertation aims to capture the sufficient evidence of these two aspects that are present in the input for the learner to generalize from. The use of expletives as a useful cue to learn raising verbs was also observed by Becker 2006, 2014, which I foreground in the analysis presented here.

In the analysis presented in this chapter, I also assume that when children hear a verb taking an infinitival complement in the input, they initially assume that it is a control construction, given the overwhelming empirical evidence in this favor mentioned earlier in this chapter (e.g., Chomsky 1969; Hirsch 2011; Hirsch, Orfitelli, and Wexler 2007; Hirsch and Wexler 2004, 2007). Another reason to assume the acquisition of control before raising stems from the fact that a majority of the verbs introduce an external argument and assign a theta-role to it; the verbs that do not introduce an external argument are much fewer in number. Children initially assume a control structure because they typically see a DP and a verb, where the DP is a thematic subject of that verb. Consequently, they assume that verbs generally introduce an external argument because they have learned this property for their language thus far. Similarly, arguments have been made suggesting that control is acquired before raising because control structures are somehow computationally simpler (Frank 1998). The hypothesis of a default control-first bias also stems from child production data, which indicates that a control structure is the default even for ambiguous verbs. The child data are shown in more detail in Section 3.4.3.

To test whether children learn from non-referential expletive subjects in the input, I examined a total of 67 control and raising predicates⁵ (Boguraev and Briscoe 1987; Postal 1974), which occur in the infinitival Pred-to-V frame in CHILDES. The predicates were identified this way as it was hypothesized that children initially identify potential con-

⁵A complete list of the verbs examined is available in Appendix A.

trol or raising verbs through a predicate’s ability to select an infinitival complement.⁶ This search allows us to first identify potential control and raising verbs. Each predicate examined occurred at least once in the corpus, ensuring that the learner would have encountered them. This resulted in 42 control, 15 raising, and 11 ambiguous predicates in total, which are in effect, an exhaustive list of these predicates that occur in child-directed speech. Although there are non-control and raising uses of these verbs, any inclusion or exclusion of the other uses of the verbs in analyzing the data is specified when necessary. The corpus search, however, ensures that each of these predicates had at least one raising or control use in the input data. Any instance of a predicate followed by an infinitival complement was considered a control or raising use of the verb. For raising verbs, instances where the predicate was followed by a *that*-clause were also considered to be verbs used in a raising construction. Any other uses were not included in raising and control-specific counts.⁷

The verbs examined were classified as either control, raising, or ambiguous primarily based on Postal 1974. Additionally, to confirm the properties of these verbs, the use of 3 tests mentioned in Landau 2013 were employed: the idiom chunks test, interpretation after passivization, and scope interpretation effects. For instance, *try* is classified as a control verb as it assigns a theta role in idioms. *Tend*, in contrast, was classified as a raising verb as it does not assign a theta role in the idiom chunks test. This is illustrated in (17).

- (17) a. #The shit tries to hit the fan. [Control]
 b. The shit tends to hit the fan. [Raising]

Similarly, there is an interpretation difference when control verbs are passivized (18); while raising verbs retain the same interpretation in both the passive and the active con-

⁶Control into purpose clauses like *John ran to avoid Mary from confronting him* were not included, as their occurrence is more widespread. I do not delve into how children distinguish between purpose clauses and other types of control constructions, as this question is not directly related to the arguments made here. The inclusion of purposes clauses only bolsters the claim made in this chapter as it would result in more evidence for control in the input. Verbs that only occur in purpose clauses were also not included in our count of control verbs as it is possible that children do not group all of these verbs in the same class.

⁷It was ensured that the criteria were met by going through the data by hand.

struction (19).

- (18) a. A specialist tried to examine John.
b. John tried to be examined by a specialist. [Control]
- (19) a. A specialist tends to examine John.
b. John tends to be examined by a specialist. [Raising]

Moreover, as expected for control verbs, *try* cannot take scope below the verb, whereas this scope ambiguity is possible with *tend*. (20) shows that with *tend*, the reading where it tends to be the case that seven students attend Glee Club is available.

- (20) a. Seven students tried to attend Glee Club. [Control]
b. Seven students tend to attend Glee Club. [Raising]

When examining the data, I also noted whether the control or raising predicate used was attested with either an inanimate or a non-referential expletive subject. This was done in order to analyze whether inanimate subjects or non-referential expletive subjects would be a more reliable indicator of a raising structure.

7 out of the 15 pure raising predicates, plus 4 ambiguous predicates, occurred with a non-referential subject. Each of the verbs identified with an expletive subject are raising verbs. These results indicate a low frequency of non-referential subjects in the data, which means that children are expected to acquire raising structures late overall, and learn the raising properties of each verb that can take an infinitival complement individually when it occurs with a non-referential subject; i.e., no generalization for raising verbs as a class is made given the input data. Only 11 out of the 25 raising predicates examined would, therefore, be learned as raising predicates in this dataset. Below are some examples of non-referential *there* occurring in subject position in child-directed speech.

- (21) a. suddenly there appeared before her the most beautiful lady she had ever seen.
(HSLLD corpus)

- b. there's gotta be a daddy shark. (Valian corpus)
- c. there happens to be a can right here. (Weist corpus)
- d. and there seems to be an awful lot less competition between them. (Cornell corpus)

English, in addition, has non-referential *it*, which also occurs with raising verbs. Some of the examples with non-referential *it* occurring in subject position in child-directed speech are provided in (22).

- (22)
- a. it started to rain (Clark corpus)
 - b. I think it's going to be okay now Ross. (MacWhinney corpus)
 - c. it's supposed to get markedly cooler. (MacWhinney corpus)

In examining the data, I went through each instance of *it* in subject position manually to make sure that it was indeed non-referential. I also used wh-questions as a test to confirm the non-referentiality of *it*; for example, the following question-answer pair for (22a) is infelicitous:

- (23) Q: What started to rain?
 A: #It started to rain.

The results from child-directed speech presented above show that non-referential subjects are present in the data as an unambiguous cue for the learner in identifying a raising verb. Since this cue is unambiguous, in principle, only one instance of a non-referential subject should suffice in indicating to the learner that the verb can occur in a raising structure. Furthermore, there is work indicating that children are aware of the presence of non-referential subjects (Chen, Valian, and Chodrow 2016; Wang et al. 1992). For instance, Chen et al. 2016 finds that English speaking children as young as 1;9 show evidence of having learned non-referential subjects when tested using an elicitation task. The CHILDES database also provides several examples of children producing non-referential subjects.

- (24)
- a. there is no moon (Valian corpus, 1;9)
 - b. there's no song like that (Naima, Providence corpus, 1;9)
 - c. its dark outside (Marjorie, Bliss corpus, 2;3)
 - d. it's cold out there (Peter, Bloom corpus, 2;5)
 - e. there needs to be another one of (Naima, Providence corpus, 3;1)
 - f. there seems to be dust here (Braunwald corpus, 3;2)

Thus, non-referential expletive subjects in English are a useful cue for learning raising predicates in English. Although relatively infrequent in the data, expletive subjects still present a reliable cue.

Moreover, none of the control predicates occur with a non-referential subject, but they do occur with inanimate subjects. The examples below show some instances where control predicates occur with inanimate subjects in child-directed speech.

- (25)
- a. cause then there will be a lot of me talking, and the tape wants to hear you talk. (Weist corpus)
 - b. a bad guys sword tried to catch you. (MacWhinney corpus)
 - c. and the foolish freight cars refused to back up. (Providence corpus)
 - d. the medicine helps to make it better. (Braunwald corpus)
 - e. the school bus decided to knock off a few kids. (Hall corpus)
 - f. really hmmm mmmhmmm what else does the fireplace like to eat (Providence corpus)

These examples illustrate that, confirming previous analyses (e.g., Becker 2006), animacy is only a statistical tendency. The data indicate that it is not a trivial matter to assume that animacy is a reliable enough cue to distinguish raising from control, especially in light of the sample size necessary given the statistical conditions.

Expletive subjects, in contrast, are a highly accurate cue. Each of the 11 Pred-to-V predicates identified belong to the raising class. The child does not encounter any false

positives, which shows that if they were using expletive subjects as cue, they would arrive at the adult grammar. Furthermore, we know that children are good at learning non-referential expletive subjects (Chen, Valian, and Chodrow 2016; Wang, Lillo-Martin, Best, and Levitt 1992). Even though the formal requirement of using a subject consistent with the target language is not in place until later stages of development, children are still aware of the referential content of the subjects even when they are missing. For instance, Wang, Lillo-Martin, Best, and Levitt 1992 found that English-learning children were able to produce non-referential subjects in an elicitation task, but none of the Chinese-learning children tested with the same task produce an overt lexical form in those instances; i.e., they were aware that non-referential subjects in their language are unpronounced, and did not use an ungrammatical overt form in those cases when prompted. Overall, we see that children are not ignoring these cues that are available to them in the input. Therefore, the use of non-referential expletive subjects is not only possible in theory, but also in practice.

2.4.2. The Lack of Overgeneralization

We have seen so far how expletive subjects allow the learner to identify a raising predicate from a purely control one; however, the question of how children learn that some predicates can be purely control without being optionally raising remains open. This question of how a learner knows that some predicates do not have an optional raising counterpart is what lead us to the learnability puzzle in the first place. We begin to answer this question by turning to the Sufficiency Principle (Yang 2005, 2016), which provides a formal model of rule learning and making generalizations. Using the Sufficiency Principle, I show that the problem of the overgeneralization of ambiguous predicates never arises. The Sufficiency Principle (Yang 2016), first stated in Chapter 1 and repeated in (26) below, allows us to calculate the number of positive members required to generalize a rule.

- (26) “The Sufficiency Principle: Let R be a generalization over N items, of which M items are attested to follow R . R can be extended to all N items iff:

$$N - M < \frac{N}{\ln N} \text{ where } \frac{N}{\ln N} := N/\ln N$$

The Sufficiency Principle captures the intuition that a large amount of evidence is needed in order for a rule to be productive. It also captures the idea of accessing information as efficiently as possible by minimizing the retrieval times of the stored items. Building on the fact that language follows Zipf's law, the Sufficiency Principle determines the point at which it is more efficient to list each item and its associated properties as opposed to forming a rule and listing exceptions in accordance with the Elsewhere Principle. The Sufficiency Principle is a psychologically real model of rule learning, and therefore, using it to determine when a rule can be generalized is desirable.

Yang 2016 provides a linguistic application of the Sufficiency Principle in addressing Baker's paradox in the acquisition of the dative alternation (Baker 1979). Yang 2016 shows that the problem of overgeneralization does arise in the acquisition of the dative alternation. Children overgeneralize the double object construction to produce sentences like *she said me no* from which they must eventually retreat. In this case, the earlier stages of acquisition reveal that the number of verbs that show the dative alternation are sufficient for this rule to be generalizable. Retreat from this overgeneralization then takes place when the number of verbs that show the dative alternation fall below the threshold required by the Sufficiency Principle; i.e., the data are insufficient for the rule to be generalizable.

Similarly, out of the 67 predicates considered here, there exists evidence for only 10 as raising. The Sufficiency Principle ($N/\ln N$) requires 51 predicates following the rule to generalize over a class of 67. Given the data, it cannot be concluded that the 11 members will be generalized to the entire class of raising and control predicates. Moreover, a non-referential subject was observed for only 4 of the ambiguous predicates. Consequently, we find that children never get to the stage where the number of ambiguous or raising predicates in the input are enough to be generalizable to the entire set of verbs. If 51 of the predicates have non-referential subjects, then the child would be tempted to assume the reasoning that the 16, which do not appear with non-referential subjects, are also raising

predicates. However, we are nowhere close to the sufficient level of positive evidence to generalize so broadly.

Table 2.5 shows verbs sorted by frequency in the child production data in CHILDES, and lists the number of raising and ambiguous predicates at the different stages. I examine the possibility of the learner either generalizing the raising property of verbs, which includes ambiguous predicates, or the learner generalizing the fact that the verbs are both raising and control. By arranging the verbs according to frequency in production, the order in which children learn these raising and control verbs can be approximated. As we can see, at no point in learning these predicates are there more raising or ambiguously raising verbs than control verbs.

Table 2.5: Control, raising, and ambiguous verbs according to their frequency in the child production data in CHILDES. The threshold of generalization is calculated using the Sufficiency Principle. Evidence for the raising or ambiguous predicate rule is never sufficient enough to be generalizable.

Verbs	Raising	Ambiguous	All Raising Productive?	Ambiguous Productive?
10 most frequent	2	2	4/10 NO	2/10 NO
20 most frequent	6	3	9/20 NO	3/20 NO
30 most frequent	8	4	12/30 NO	4/30 NO
40 most frequent	10	6	16/40 NO	6/40 NO
50 most frequent	12	9	21/50 NO	9/50 NO
60 most frequent	13	11	24/60 NO	11/60 NO

As the table above indicates, the learner does not have sufficient evidence to generalize the raising rule to the entire class of predicates at any point in learning these predicates. Therefore, the learner will have to lexically learn the predicates that are raising on an item-by-item basis. We also see from the fact that there is only positive evidence for raising for 11 verbs out of 67, that the control structure can be overgeneralized. The

Sufficiency Principle requires at least 16 exceptions to a class of 67 for the majority rule to be unproductive. In this case, with only evidence for 11 verbs as raising, the learner overgeneralizes the control structure for the 66 verbs. Since more evidence for the control structure of verbs is present in the input, the learner has a control-predicate default.

Since the number of ambiguous or pure raising predicates is not large enough for the learner to overgeneralize, the learnability problem does not arise. The learner only overgeneralizes if there are less than $N/\ln(N)$ exceptions to a rule, and here, the number of exceptions well exceeds the number of predicates that show properties of raising. This mistreatment of the problem of generalization in the first place resulted in the learnability puzzle, which does not exist.

Now that we have discussed distinguishing the raising predicates from control ones, let us turn to how the learner identifies ambiguous predicates from the pure raising ones in the following section.

2.4.3. Ambiguous vs Pure Raising Predicates

Out of the 67 predicates examined, 11 belong to both the raising and control class. These verbs were *fail*, *begin*, *continue*, *manage*, *need*, *have*, *promise*, *stop*, *grow*, and *start*. In this section, I show that one difference between ambiguous and pure raising predicates that can be exploited by the learner is the rate of expletive subjects they occur with. An ambiguous predicate should, in theory, occur with expletive subjects at a lower rate than pure raising verbs, as they have an optional control counterpart. This distinction between the class of ambiguous and pure raising predicates can easily be captured and quantified under a competition approach; i.e., under the variational model (Yang 2002, 2004), which I adopt here. The variational model encodes competition between two options such as two grammars, lexical items, or verb classes. The model is briefly described below before it is applied to the case at hand.

Under the variational model, the child selects a particular grammar with the proba-

bility determined by the evidence in the linguistic input. A grammar is punished if the input is incompatible, and rewarded if the data is in line with it.⁸ The variational learning model is different from the Sufficiency Principle in that it makes reference to the token frequencies of forms, whereas the Sufficiency Principle uses the type frequencies within lexical classes. Another crucial feature of this model is that a grammar, or a parameter value, can potentially never be fully eliminated. Therefore, it is possible for two values of a parameter to be stored.

The variational model is distinct from Becker's model in a some respects. One crucial difference between Becker's model and the variational model is that Becker's model evaluates the data globally at a certain point. The variational model, on the other hand, incrementally evaluates the different competing forms. A second difference between the two models is that the variational model does not compare the rates of raising and control verbs.

I use the variational model here to distinguish between verbs that store their control counterpart to varying degrees. First, we can show that indeed, pure raising verbs and ambiguous predicates occur with non-referential expletive subjects at different rates. I investigated ambiguous and pure raising predicates in the input data, and found that ambiguous predicates occur with expletive subjects at a lower rate than pure raising predicates. A pure raising verb like *seem* occurs with far more expletive subjects than an ambiguous verb like *begin*. As a result, *seem* more readily yields to raising properties than *begin*. The analysis of the acquisition of pure raising verbs versus ambiguous verbs is analogous to other word learning models, such as the pursuit model of word learning (Stevens, Gleitman, Trueswell, and Yang 2017). Under the analysis described here, the competitor control meaning of a verb is ruled out with enough instances of non-referential subjects in the input.

The relative frequencies of non-referential subjects in the input can be captured by the

⁸See Yang 2002 for further details on the variational model, and Legate and Yang 2007 for an example of implementing the variational model.

variational model (Yang 2002), which allows for competition between two forms. Here, both the raising and control forms of a verb, which are learned independently, can compete with each other depending on the data in the input. The quantity of the input experience determines whether the raising form wins out over the control form, or if both forms of the verb are stored. For the ambiguous predicates, the learner learns a probabilistic distribution over two forms: raising and control. Hearing a non-referential subject in the input points the learner towards the direction that the predicate is raising. A referential subject does not necessarily indicate that the verb assigns a theta role, as the theta role could have been assigned to the referential DP by a verb in a lower clause. Only referential subjects in monoclausal structures can provide such a cue.

To test whether ambiguous predicates did indeed occur less frequently with non-referential subjects, as opposed to pure raising predicates, I conducted a corpus search on the collective data for all English learning children in CHILDES. The test case used was that of *begin* versus *seem*, which occurred a comparable number of times in the input, at 464 and 468 total instances, respectively. The results of the corpus search show that *begin*, an ambiguous predicate, occurred 5 times only with non-referential *it*, but did not occur with *there*. *Seem*, on the other hand, occurred with non-referential *it* 13 times, and 9 times with *there*. These numbers are summarized below in Table 2.6.

Table 2.6: Non-referential subjects with *seem* and *begin* in control constructions in the input data from CHILDES. The rate of non-referential expletive subjects with *seem* and *begin* in control and raising constructions are significantly different ($p < 0.01$).

	<i>seem</i>	<i>begin</i>
Occurrences in control and raising constructions	266	130
Occurrences with non-referential <i>it</i>	13	5
Occurrences with <i>there</i>	9	0

The total number of non-referential expletive subjects noted for *seem* in raising constructions in CHILDES was 22, while there were only 5 instances of non-referential subjects

with *begin*. The results from CHILDES suggest that the two predicates can be differentiated on the basis of the use of non-referential subject. It may be numerically unclear whether the proportions of non-referential subjects with the two verbs are different enough from each other. Therefore, I conducted a two-sided chi-squared test with continuity correction on the data, which revealed that the two proportions are significantly different ($p < 0.01$). Table 2.7 takes into account all occurrences of *seem* and *begin*, which include non-raising and control constructions, and shows that even though the number of expletive subjects with *begin* remain the same, the number of non-referential *it* subjects with *seem* is much higher. In all occurrences of *seem*, 48 non-referential *it* subjects were found, resulting in a total of 57 non-referential subjects. The difference in the rate of expletive subjects between the two kinds of predicates is even more pronounced when taking all construction types into consideration. To confirm that the two proportions were significantly different, a two-sided chi-squared test with continuity correction was conducted. The test confirms that the rate of inanimate subjects with the two predicates, *seem* and *begin*, are indeed different ($p < 0.01$).

Table 2.7: Non-referential subjects with *seem* and *begin* in all constructions in the input data from CHILDES. The rate of non-referential expletive subjects with *seem* and *begin* is significantly different ($p < 0.01$).

	<i>seem</i>	<i>begin</i>
Total occurrences	464	468
Occurrences with non-referential <i>it</i>	48	5
Occurrences with <i>there</i>	9	0

Thus far, we have seen that the data are compatible with an analysis where children use non-referential expletive subjects to differentiate between pure raising and ambiguous verbs. We can now also further test whether the data are compatible with an animacy-based approach. When investigating raising and control constructions, in contrast to what we find with expletive subjects, the rate of inanimate subjects between *seem* and *begin* is

not significantly different. We find that in raising and control constructions, *seem* occurs with 82 inanimate subjects, and *begin* occurs with 30 inanimate subjects, including expletives for both. These numbers are summarized in Table 2.8. Thus, the rate of inanimate subjects with *seem*, a raising predicate, is 0.3 and the rate of inanimate subjects with *begin*, an ambiguous predicate, is 0.23. It should be noted here that Becker 2014 found that *seem* occurs with a higher number of inanimate subjects than other raising verbs, although *begin* was not found with the same number of inanimate subjects in her case. To test whether the proportion of inanimate subjects was significantly different, a two-sided chi-squared test with continuity corrections was conducted. The results show that even though both *seem* and *begin* occur with inanimate subjects, the difference between the rates of the inanimate subject is not significant ($p = 0.14$).

Table 2.8: Inanimate subjects with *seem* and *begin* in control constructions in the input data in CHILDES. The rate of inanimate subjects with *seem* and *begin* in control and raising constructions are not significantly different ($p = 0.14$).

	<i>seem</i>	<i>begin</i>
Occurrences in control and raising constructions	266	130
Occurrences with inanimate subjects	82	30

To corroborate what was shown earlier using the rates of inanimate subjects with raising and ambiguous verbs from Becker 2014’s own calculations, I also show the number of instances needed to distinguish between pure raising and ambiguous verbs using our data, which includes all of the input data in CHILDES. When taking into account the rate of inanimate subjects with *seem* and *begin* in all constructions, we find that *seem* occurs with 106 inanimate subjects, and *begin* occurs with 123 inanimate subjects. The rate of inanimate subjects between these two predicates remains not significant when considering all occurrences of the two predicates as well ($p = 0.3$).⁹

⁹*Begin* and *seem* are good test cases for the indirect negative evidence approach as they occur more frequently in the data and with more inanimate subjects when compared to other raising or ambiguous

Table 2.9: Inanimate subjects with *seem* and *begin* in all constructions in CHILDES. The rate of inanimate subjects between the verbs is not significantly different ($p = 0.24$).

	<i>seem</i>	<i>begin</i>
Total occurrences	464	468
Occurrences with non-referential <i>it</i>	106	123

Moreover, performing the same calculations as we did earlier in Section 3.2.3, we find that about 600 instances of the predicates are required to reliably distinguish between the two classes with a 95% confidence interval. This is seen in Figure 2.5, which illustrates the number of trials needed to distinguish between two predicates of type A and B with rates of observing inanimate subjects at 0.23 and 0.3 respectively. With a 99% confidence about 720 instances of the predicate must be observed before the learner can determine its class with any certainty. The number of encounters needed for these verbs is more than their total occurrences in all the input in the CHILDES database combined. This further indicates that learning from animacy under an indirect negative evidence approach requires more evidence than may be available to the learner. Thus, the difference in the rates of non-referential subjects can be used to distinguish between ambiguous and pure raising predicates, but the rate of inanimate subjects between the two classes cannot be used similarly.

Additionally, in order to control for the relatively small size of the CHILDES corpus, I examined the occurrences of *seem* and *begin* in the Corpus of Contemporary American English (COCA). This larger corpus allows us to better gauge the data available in the input using the total of over 520 million words of text available. In COCA, we find that *seem* and *begin* occurred about 278,531 and 234,357 times respectively. As we saw in CHILDES, the token frequency for these predicates is roughly the same, which is precisely why they were chosen to test the relative frequencies of non-referential subjects. The corpus

predicates. For instance, *tend* occurs only 15 times in a raising or control construction in the data with 4 inanimate subjects, and *sure* only occurs 26 times in the data with 2 inanimate (expletive) subjects.

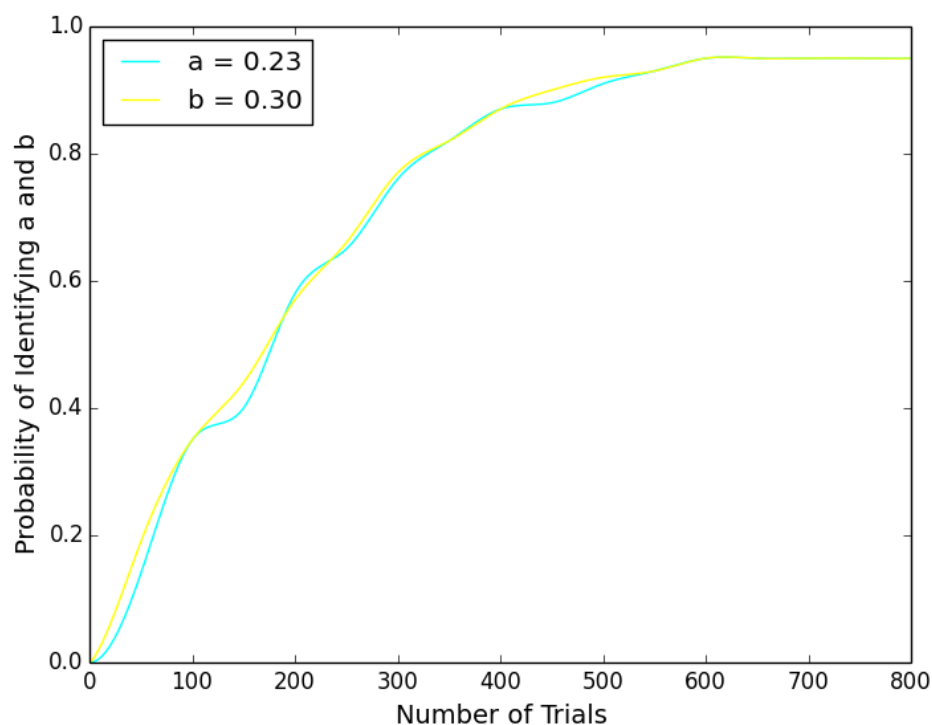


Figure 2.5: Number of occurrences needed to distinguish pure raising predicates and ambiguous predicates with a 95% confidence interval when using the rate of inanimate subjects from *seem* and *begin* in child-directed speech.

was searched for *there*, and found that while *seem* occurred with *there* 3,514 times, *there* appeared with *begin* only 119 times.¹⁰ These results are summarized in Table 2.10.

Table 2.10: Expletive Subjects with *seem* and *begin* in COCA. The proportion of non-referential expletive subjects with *seem* and *begin* are significantly different ($p < 0.01$).

	<i>seem</i>	<i>begin</i>
Total occurrences	278,531	234,357
Occurrences with <i>there</i>	3,514	119

The results of 119 versus 3,514 show that there is a substantial difference between the number of non-referential subjects that occur with ambiguous predicates and those that

¹⁰Since examples with non-referential *it* need to be combed through by hand, they were left out here due to the size of the data. For the same reason, we do not provide the rate of inanimate subjects with *seem* and *begin* in COCA.

occur with pure raising ones. A two-sided chi-squared test with continuity correction also shows that the two proportions are significantly different ($p < 0.01$).

Under the variational model, the data found are compatible with the predictions for whether the control or raising forms of the verbs are stored. A child has more evidence for the raising counterpart of *seem*, than for *begin*. This was determined through the significant difference in the rate of non-referential subjects of the two verbs. Thus, both control and raising forms for *begin* are stored, whereas the raising form for *seem* wins out and its control form is lost; i.e., 3,514 occurrences of a non-referential subject is enough for the learner to conclude that *seem* is a raising predicate, whereas 119 occurrences of a non-referential subject leads the learner to assume that *begin* is ambiguous between having a raising and control structure.

The idea that *begin* and *seem* are learned in relation to the number of non-referential expletive subject uses finds support in children's production data in CHILDES. Although *begin* and *seem* occur with the same token frequency in the input, the number of non-referential subjects with which the predicates are used differ. The significantly different rates of non-referential subjects with ambiguous and pure raising predicates can be exploited by the learner. Child production data suggests that children are sensitive to this distinction. A search through the combined child data in CHILDES returned no uses of *begin* with a non-referential subject. In contrast, *seem* is used with expletive subjects at various ages. A few examples of *seem* used with an expletive subject are shown below:

- (27) a. there seems to be dust here (Braunwald corpus, 3;2)
b. it seems like this goes here (EllisWeismer corpus, 3;7)
c. it seems to me that's all Graeme is taping today (Hall corpus, Jub, 4;9)
d. Mom it doesn't seem like he's green (HSLLD corpus, 5;8)
e. it seemed like they didn't know where they [/] they were either (Gillam corpus, 6;2)

The data above show that children are producing expletives with *seem*, and therefore, they are at least aware of the raising property of *seem*.

In addition to the positive evidence seen for the raising form of a verb, there are also constructions that positively indicate that a verb is a control verb. When the learner hears a monoclausal sentence with a DP object as in *The CEO began the meeting*, it is clear that the object originates within the same clause and does not originate from a lower embedded clause. In this case, the DPs both have their theta-role assigned by the verb *begin*, and *begin* assigns case to its object. This is also seen with control verbs like *hope* that do not take DP complements. In a sentence like *John hoped for apples*, there is no lower clause that the DP could have originated from. Pure raising verbs are never found in this structure. Such simple transitives are frequent in the data even with ambiguous predicates like *begin*. Some examples from child-directed speech are provided below.

- (28) a. Where together they began a grand tour of the entire art museum. (Providence corpus)
- b. They began their journey toward the surface. (Providence corpus)
- c. After a little while, the magic carpet began its descent. (Providence corpus)
- d. I've started my diet today. (Hall corpus)
- e. I think she started it. (Hall corpus)
- f. We just started a new sticker chart. (Weist corpus)
- g. Did we grow these outside this summer? (Valian corpus)
- h. They didn't grow them. (MacWhinney corpus)

Not only is this cue present in the input, but children are aware of it as well. Examining the child production data reveals children using simple transitive sentence of this kind with ambiguous verbs, and never with pure raising verbs. Some examples are provided in (29).

- (29) a. he began the walk here (Sachs corpus, Naomi, 3;5)

- b. we don't grow em (NH corpus, Edwina, 4;0)
- c. this is a picture where they grow corn in Texas (Kuczaj corpus, Abe, 4;10)
- d. but then later on after the show they grew them (MacWhinney corpus, Ross, 5;4)
- e. I need it (Braunwald corpus, Laura, 2;2)
- f. I need the other baby (Davis corpus, Cameron, 2;8)

The examples above show that positive evidence for verbs that take a control structure is also present in the input. By using expletives as a positive cue for raising together with the kind of data in (28), the learner acquires pure raising verbs, and verbs that are ambiguous between control and raising.

To summarize what we have seen thus far, Becker 2006, 2014 propose one learning model for the three kinds of predicates, where the learner chooses one out of the three verb classes in consultation with the probabilities of the verbs occurring with an inanimate subject. Instead, I have argued that the learner only starts off by assuming a default control structure. When the verb occurs with non-referential subjects, the learner assumes it's a raising verb. Once the verb has been postulated to be both raising and control, then the two uses are in competition. Then, the learner must distinguish pure raising verbs from ambiguous verbs.

Throughout this chapter, I have argued that non-referential subjects are a reliable and sufficient cue to the learner to identify raising predicates. The calculations performed above also indicate that even if we were to put animacy in the variational model, it might not work to distinguish between the classes of predicates. Ambiguous predicates are learned based on positive evidence in the input for raising, and positive evidence in the input indicating that a verb assigns a theta-role to its external argument, a defining property of control verbs. That said, it should also be noted that there may be other tendencies present in the input as well, such as the tendency for more inanimate subjects to occur with raising verbs. However, the data presented in this chapter suggest that the rate of

inanimate subjects cannot be taken to license categorical generalizations of predicate type, even though the tendency of the predicates to take inanimate subjects at different rates may be helpful to the learner.

2.5. Conclusion

This chapter shows that the overwhelming majority of the data are consistent with the child analyzing all potential predicates as being control initially, possibly as a default strategy or due to the simplicity of the grammatical structure (e.g., Wexler 1992). The data also show that the non-referential subject-driven strategy fares better than the animacy-based model, and moreover, learning ambiguous predicates is not problematic under this analysis. For an indirect negative evidence learner, the data in the input are noisy, and the distinction between the two classes is not unequivocal. In contrast, one would not have to rely on heavy evidence in the input when dealing with positive evidence. In the case of positive evidence, every datapoint in the input is accurate. A predicate is identified as raising, iff it appears with a non-referential subject. The findings presented here are consistent with previous work that argues for the acquisition of control before raising.

Another core issue at stake here is the problem of generalization. The supposed learnability problem discussed in this chapter resulted from the misuse of the problem of generalization (Becker 2006). The symptom of such a mistreatment is the kind of argument that if in some case, even one member in a class shows a certain property, then the learner will need to rule out the likely superset hypothesis that all members in that class show that property, even when many other members in that class do not. In this instance, “some members” are verbs like *begin*. The problem here is taken to be that, the child would consider all control verbs as candidates for raising, even though only a few verbs show this property.

I have shown that the Sufficiency Principle, which is adopted here, provides a general

solution to the problem of when to generalize. If there are more than $N/\ln(N)$ exceptions for N potential members within a class, then the rule is not productive and the learner does not generalize. Only if a sufficiently large number of members in a class show X in the input data, then the child generalizes to the entire class. Otherwise, X is lexicalized for those specific members. Generalization only occurs when the number of exceptions to a class with N members that could follow rule R is less than $N/\ln(N)$. In the present case, I showed that there are not enough verbs such as *begin* that meet the threshold; therefore, the child does not generalize. The problem of generalization never occurs. In fact, the analysis proposed here can also be extended to claim that we would never expect to see a language where raising predicates would be overgeneralized. In addition, the child also only uses unambiguous cues— i.e., non referential subjects— to identify raising verbs.

Chapter 3

The Acquisition of Causatives

3.1. Introduction

In English, some verbs, like *break*, can occur both as an intransitive (e.g., *the vase broke*) or as a transitive causative (e.g., *John broke the vase*). Many verbs display this property, known as the causative alternation, but not all verbs that occur as an intransitive have a transitive causative counterpart (e.g., *fall*, *disappear*). Adults are well-aware of this restriction, but children, on the other hand, make overgeneralization errors (Bowerman 1982; Pinker 1989) by extending the causative form to intransitive only verbs. Some of these errors are shown in (30).

(30) Child causative errors (from Bowerman 1982):

- a. He's gonna die you David (4+)
- b. Kendall fall that toy (2;3)
- c. You ached me (4;1)
- d. She came it over there (3;4)

These errors indicate that there is a time when the learner has acquired a grammar that is a superset of the adult grammar. The errors involve the learner extending a causative form to verbs that could potentially be used as lexical causatives, but are not. For instance, in English, it is possible to say *John broke the vase* to mean that John did something that caused the vase to be broken, but **John fell the vase* is ungrammatical even if John directly

did something that caused the vase to fall. Adults adhere to these grammaticality patterns, but children, as we have seen, overgeneralize to produce these errors. Since the data provided by Bowerman 1982 are not available in the public domain, I provide examples of some additional causative errors found in CHILDES below.

(31) Causative errors in CHILDES:

- a. Please stay her (Bohannon corpus, Nathaniel, 3;00)
- b. I'm want them to disappear them again (MacWhinney corpus, Ross, 3;02)
- c. I wanna go it this way or this way (Kuczaj corpus, Abe, 03;04)
- d. Hey we're gonna fall it (Hall corpus, Steven, 4;09)

The examples provided in (31) illustrate that children robustly causativize, and as a result, make overgeneralization errors. The errors indicate that children take verbs that occur as intransitives, and use them in a causative transitive frame. In this chapter, I address the question of how children come to generalize the causative alternation rule. Moreover, children at this stage must eventually retreat from their overgeneralization in order to arrive at the adult grammar, which is a subset of their generalized superset grammar. Previous research has suggested either strict semantic bootstrapping accounts or learning from indirect negative evidence as a solution to the problem. Instead, I argue that the causative alternation rule is learned from positive evidence in the linguistic input.

The rest of this chapter is structured as follows. First, I describe previous work on the acquisition of causative verbs and the causative alternation construction. Next, I evaluate accounts that propose that syntactic structure is acquired from verb semantics (Pinker 1989), as well as usage-based accounts. In Section 3.4, I present a novel solution to the problem; learning from positive evidence. In doing so, I address the conceptual and structural cues that the learner uses in the path to arriving at the adult grammar. I show how the developmental trajectory of the learner, which consists of the child's overgeneralization and retreat from overgeneralization, can be accounted for using the Active Mapping

Model.

3.2. Background

This section provides a background on the kind of causative errors observed in child speech. I also discuss previous approaches that have been proposed to account for these errors. I describe two lines of work: one that appeals to indirect negative evidence in the form of statistical preemption and entrenchment as learning mechanisms (e.g., Ambridge, Pine, Rowland, and Young 2008; Bowerman and Croft 2008), and the other that proposes linking rules between verb semantics and their argument structure (e.g., Pinker 1989). After describing these approaches, I show how neither line of work satisfactorily captures the developmental facts of the acquisition of the causative overgeneralization.

Before diving into previous accounts, Sections 3.2.1 and 3.2.2 first describe the nature of child causative errors and the nature of causative alternation rule in the adult grammar. I propose an analysis of how the learner transitions from the child grammar to the end state, the adult grammar, by investigating both the adult grammar and the errors made by the learner in more detail. In doing so, I show that children generalize the causative rule to pure unaccusative verbs. The sections to follow also illustrate that even though not all unaccusative verbs causativize in the adult grammar, the causative alternation rule is productive for certain subclasses of verbs.

3.2.1. Description of Causative Errors

This section describes the causative errors made by children in more detail. The causative errors observed by Bowerman 1982 occur with verbs belonging to a wide range of semantic subclasses. However, at the same time, there are also similarities between the kinds of verbs that the learner erroneously causativizes. This point is crucial as children extend the transitive causative form to a particular type of intransitive verb, while still using the

causative form on intransitive verbs across a diverse range of meanings. The verbs and their association with a narrow semantic class, as described by Pinker 1989, is provided below. The following table of overgeneralized verbs is from Pinker 1989, who summarizes the errors made by Bowerman’s children.

Table 3.1: Summary of verbs overgeneralized in a causative frame by the children observed by Bowerman 1982. The semantic subclass and number of errors noted with each type of verb is also provided (Pinker 1989:303, Table 7.8)

Subclass	Verbs	# of Sentences
Directed motion	come, go, fall, rise, drop	30 (28%)
Going out of existence	die, disappear, vanish	12 (11%)
Being/staying	stay, be, spell, sound, wait	16 (15%)
Possession	have, take	13 (12%)
Psychological	remember, watch, guess, wish, feel, ache, learn	12 (11%)
Involuntary emission	sweat, blood	3 (3%)
Internally caused state change	bloom	1 (1%)
Semivoluntary expression of emotion	laugh, cry, giggle	5 (5%)
Voluntary action	eat, drink, sing, talk swim, climb	14 (13%)

Table 3.1 provides an overview of the kind of verbs that children erroneously causativize.¹ As can be seen from this table, these verbs belong to a variety of semantic subclasses, but the property they have in common is that all of these verbs lack intentionality. At first glance, it appears that all but one subtype of errors, verbs of voluntary action, belong to unintentional verbs; however, upon closer look it appears that the voluntary action subclass of verbs is only used causatively in cases when the object is inanimate. Some

¹The work on acquisition done to date primarily refers to these original set of errors (e.g., Ambridge, Pine, Rowland, and Young 2008). However, the production data from these children are not available in the public domain, and therefore, the analysis presented in this chapter will largely be based on child data available in CHILDES. These errors are primarily intended to provide an overview of the type of verbs that children make causative errors with.

examples of causative errors with these “voluntary action” verbs are shown below.

(32) Examples from Bowerman 1982 as cited by Pinker 1989:

- a. I wanta swim that [Holding an object in the air and wiggling it as if it were swimming] (E, 2;1)
- b. I’m talking my birdie [Pulling string on bird-shaped music box] (E, 2;2)
- c. I’m singing him [Pulling string on cow-shaped music box] (C, 3;1)

The examples in (32) are compatible with the idea that the learner is aware of the distinction between intentional and unintentional verbs. In these cases, it is likely that the learner is still following the causativization rule for unintentional verbs, while treating certain typically intentional verbs as unintentional given the circumstance. Foreshadowing the analysis presented in this chapter, I argue that children use intentionality in acquiring verb classes in their language. I return to this point in later sections.

3.2.2. Experimental Evidence for the Productivity of the Causative Alternation

In the previous section, we saw that children produce causative errors with verbs that generally lack intentionality. Up to this point, we have also primarily focused on the fact that children generalize the transitive causative form to verbs that do not occur in that frame in the adult grammar. However, an important aspect of the causative alternation is that it is productive in the adult grammar for certain semantic subtypes of verbs. We focus on this point in this section.

This section summarizes previous experimental results that find that intransitive (unaccusative) verbs are productively causativized when they involve external causation in the adult grammar. The productivity of the causative alternation is important because any analysis of the acquisition of causatives must also account for the productivity of this alternation. A strict lexicalist hypothesis, for instance, cannot be correct here.

As we have seen in previous sections, studies on the causative alternation have often emphasized that children use the causative rule productively. To supplement these findings, Maratsos, Gudeman, Gerard-Ngo, and DeHart 1987 tested whether children and adults would generalize novel intransitive verbs as a transitive causative. In an experimental setting, they introduced a novel intransitive verb *fud*, and examined the spontaneous speech of adults and children when they used this novel verb in the testing phase of the experiment. The experiments were conducted on adults and on children aged 4;6 – 6;2. The novel verb *fud* in the causative referred to a dough-like substance being converted into strands by the means of a machine. In the intransitive, it referred to a dough or clay-like state. In the experiment, a total of six sessions with story-telling and demonstration tasks for spontaneous production were used overall, in addition to a sentence completion task. 40 children in total, who were assigned to four groups, participated in the experiment. Group I heard *fud* only as a transitive (control group), while group II heard *fud* only in an intransitive frame. Those in group III were exposed to sentences that were intransitive with a locative or benefactive, and Group IV also heard *fud* as an intransitive, but the children here were not encouraged to use the verb overtly in any way; they were simply told that they were going to learn about *fudding*. Moreover, *fud* was never used as a periphrastic causative in the exposure phase of the experiments.

The results of the study found that children used *fud* as a verb ranging from 8 times by one child and 107 times by another. If a verb was used with an agent, it was counted as transitive even when the object was omitted. Although the results varied between groups, preference for the use of the causative form was found overall. Group I, the control group, produced transitive causative sentences 98% and 97% of the time in story and demonstration tasks when they heard the novel verb used in the transitive causative frame. Those in group II produced the causative form 23% of time in spontaneous production and 37% of the time in the sentence completion task even when they did not hear the novel verb used in a transitive frame. Group III produced *fud* 27% and 23% of the time in spontaneous

speech and the sentence completion task respectively, and group IV did so 21% and 30% of the time. Even when the children did not hear the novel verb used in a causative frame, children produced the verb transitively 26% of the time overall. The number of transitive causatives produced by each subject in groups II through IV Maratsos et al. 1987's experiment is shown in Figure 3.1. There were a total of 28 participants in these groups where the novel verb was only used as an intransitive. This figure is taken from Maratsos et al. 1987 (p.101, Figure 4.1). These results are indicative of the productivity of the causative alternation rule.

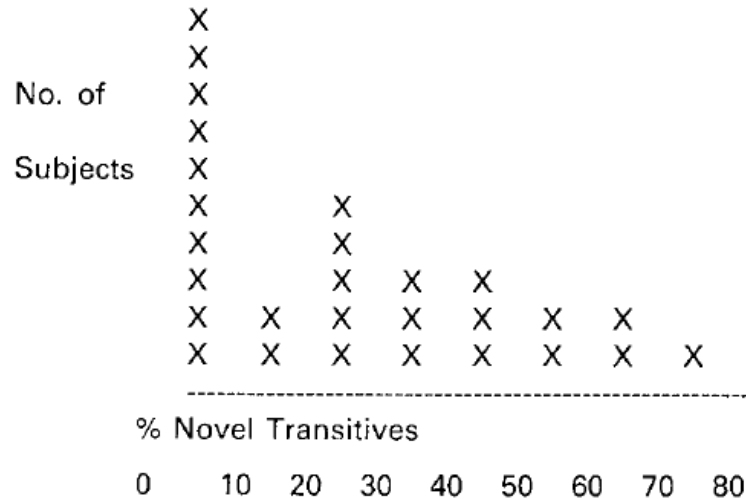


Figure 3.1: Number of transitives produced by each subject in Maratsos's et al. 1987's experiment. The results of the participants illustrated here are from the non-control groups where only the intransitive form of the novel verb was used (N = 28).

The results from Maratsos et al. 1987's study illustrate that some children are willing to causativize a novel verb, but there is also individual variation where a number of participants did not causativize the novel verb. Additionally, Gropen, Pinker, Hollander, and Goldberg 1991 also find evidence for the productivity of the causative rule in English. They find that children causativize intransitive verbs when the action depicts a direct causative relationship. Thus, these findings also support Bowerman 1974's claim

that children produce these causative errors in a systematic way.

The results from Maratsos et al. 1987's study and Gropen et al. 1991's study show that children and adults both readily and productively causativize verbs denoting direct external causation from their intransitive frame. The reverse was not found. Thus, throughout this chapter, I assume that children also causativize from a base intransitive form, and the generalization under question is that from an intransitive to a causative transitive frame.

Thus far, we have seen that the causative rule is productive when direct external causation is depicted. There is also evidence in the literature that other semantic subtypes of verbs productively undergo the causative alternation (e.g., Brooks and Tomasello 1999). For instance, Brooks and Tomasello 1999 tests two novel verbs, one with manner of motion semantics (e.g., *roll*) and the other with directed motion semantics (e.g., *come*). They found that children over the age of 4 causativized the novel manner of motion intransitive verb, but not the novel directed motion intransitive. These results indicate that children eventually learn that certain semantic subclasses of verbs can productively causativize, even if not all intransitive verbs undergo the alternation.

Throughout this section, we have seen evidence for the productivity of the causative alternation. We bear this productivity in mind as a theory that assumes a strict conservative form of lexical learning would not predict that these forms should occur productively in the first place. The results also indicate that participants were almost always unwilling to use a novel verb that they had heard in a transitive causative frame as an unaccusative without positive evidence indicating that the verb could occur as an intransitive. In the following section, I discuss lexical semantic and usage based approaches that provide an account of the productive and unproductive aspects of the causative alternation rule, and show why they are inadequate in accounting for the problem.

3.3. Lexical Semantic and Usage-based Approaches

In the following subsections, I describe two main approaches to the acquisition of the causative rule. One line of work argues for a semantic bootstrapping analysis using innate linking rules between the semantics and the argument structure of the verb that guide the learner to the adult grammar (Pinker 1989). Another approach uses indirect negative evidence in the form of statistical preemption and entrenchment analyses to account for the retreat from overgeneralization (e.g., Ambridge 2013).

Statistical preemption accounts claim that the ungrammatical causative errors can be overridden when the child learns the grammatical, suppletive causative form of the same verb. These are cases where the causative form of the inchoative verb appear to be suppletive such as *kill* and *die*. Such an approach is used by both semantic bootstrapping and usage-based accounts. In the following sections, I show that statistical preemption cannot adequately result in retreat from overgeneralization, and moreover, there appears to be no effect of preemption on retreat from causative errors (Ambridge 2008; Bowerman and Croft 2008).

Additionally, I argue against an entrenchment account, which claims that if a verb is heard in an intransitive frame enough times, the learner will assume that it can only occur as an intransitive. I provide evidence against this approach by showing that verb frequency alone cannot account for the overgeneralization and the retreat from overgeneralization.

I describe the approaches of statistical preemption and entrenchment in more detail in Sections 3.3.1 and 3.3.2. Then, in Section 3.3.3, I provide a critical assessment of these approaches and show why they do not adequately account for the problem.

3.3.1. Deriving Argument Structure from Verb Semantics

This section describes an approach to the acquisition of causatives (Pinker 1989) that proposes that children learn the argument structure of verbs from the verb semantics. Pinker 1989 argues for linking rules between the semantic structure and the argument structure of a verb. It is the semantic criteria on lexical rules that restrict the general application of the causative alternation. Under this analysis, the argument structure properties follow from the semantic properties of the verbs, which are largely constrained from the onset. In other words, under this approach, the semantics determines the argument structure of the verb. In this section, I describe two kinds of semantic analyses, a linking rules analysis (Pinker 1989) and an event semantic analysis (Ambridge et al. 2011, 2018).

I first describe Pinker 1989's approach² here in detail as it is influential in the literature on the acquisition of causative verbs, and indeed, the acquisition of verb argument structure properties overall. To account for how children generalize the causative alternation rule, and how they eventually retreat from this overgeneralization, Pinker 1989 proposes that verb argument structure is derived from the verb's semantics. The verb learning mechanism that is proposed in Pinker 1989 is that verbs have an initial semantic structure with a syntactic structure associated with it, which is largely available to the learner from the beginning. In later stages of acquisition, the learner may posit a rule that takes this semantic structure as the input and alters it. The rules, thus, apply directly to the semantic structure of verbs, which in turn are linked via linking rules to an argument structure associated with the new, altered semantic structure of verbs. When the learner acquires a verb, the verb is part of a larger verb class that shares the same semantic properties.

One consequence of this approach is that if a verb belongs to a class that shows the causative alternation, e.g., change of state verbs, then each and every verb that falls into this class must alternate. If a verb falls into a class that does not alternate, e.g., the class

²The reader is referred to Bowerman and Croft 2008 for a thorough critical assessment of Pinker 1989's analysis of the acquisition of causatives.

of directed motion verbs, then no verbs in this class can occur as a transitive causative. Thus, such an approach cannot readily tolerate exceptions to the causative alternation rule. The causative errors made by the learner indicate that the verbs that children are drawing parallels between are verbs that can alternate and verbs that do not, such as *fall* and *disappear*. A linking rules account does not predict these errors. These errors are only predicted if the learner falsely semantically classifies a verb, and therefore, under this account, these errors are not the result of a productive rule in the learner's grammar. Instead, they are classified as a mistake that is later rectified in the acquisition process.

In the case of causatives, Pinker argues that verbs with no directly causable change are inherently incompatible with the causativization rule; i.e., the semantics of these verbs are incompatible with the causative alternation. The learner does not learn these restrictions imposed on the verb, instead, the lexicon of a language specifies the semantic subclasses that guide the learner to identify the verb's semantic choosiness.

Under Pinker 1989's approach, the semantic properties of verb classes are delineated as broad range and narrow range rules. Broad range rules are properties that work top down within a class, while narrow range rules work bottom up. Pinker proposes these rules to account for the productivity of the causative alternation. These broad range rules explain the productivity of a rule. As a counterpart to broad range rules, Pinker 1989 proposes narrow range rules that are meant to restrict all verbs that meet the criteria of direct causable change. These are the verbs that fall within a large verb class with certain semantic properties, but still do not display some of the properties typical for members of that class. For causatives and causativization, for instance, verbs that undergo a change of state, such as *melt*, *open*, etc., are argued to undergo the alternation as change of state verbs typically display this property. In this case, the broad range rule is that change of state verbs alternate. However, not all verbs that denote a change of state, like *disappear*, follow this rule. Therefore, for this class, the broad range rule of change of state is a necessary, but not a sufficient condition for the alternation.

In addition to the claim that overgeneralization errors occur due to the application of broad-range rules before learning narrow range rules, Pinker 1989 also argues that child errors result from systematic misconceptions about meanings of particular verbs (Pinker 1989:292). For instance, children are argued to conflate an inherently directed motion verb like *come* with a manner of motion verb like *roll*. The overgeneralizations found result from erroneously causativize verbs like *be* and *come* because the discourse context calls for a lexical causative that the learner has not yet acquired. According to this analysis, this situation never arises for adults as there is always a high frequency grammatical causative form available to them. Pinker 1989 also cites Clark 1987 in saying that children stretch the boundaries of their knowledge as they have pervasive lexical gaps. They lack the adult means to communicate. For instance, they might not have learned the causative forms like *keep* and *bring* properly early on, which is why they make errors with the intransitive forms like *come*. It is suggested that as transitives are remastered, the causative errors decline. This argument consists of three parts: first, children's use of argument structure alternation is always semantically conditioned. Second, children's overgeneralizations are generally due to the overapplication of broad-range rules, and third, children's overgeneralizations are due to incorrect verb meanings.

As stated earlier, a relevant semantic class and the properties associated with the class, under this approach, is innately present in the child's language learning mechanism. This type of learning model constrains learners from the beginning, as the learner is posited to know the correct argument structure of the verb once they know the verb semantics. Given that many of these verbs are extremely frequent in the input (e.g., *come*, *go*), it cannot reasonably be claimed that the learner is unaware of the semantics of these frequent verbs. The fact that the verb argument structure is available to the learner for free once the verb semantics are acquired, the linking rules model does not predict that the learner would make argument structure errors. The innate syntax-semantic mapping of the linking rules model has the effect of relegating the learner's causative errors to the

status of simple mistakes that do not reflect any underlying generalization. There cannot be a productive rule for the learner's causative errors, as the semantic subclass that the verb is associated with either shows the causative alternation or the class is a class where the verbs are pure intransitives. Therefore, the observed child errors in this case can only be due to inaccurate learning of the verb semantics. Under this account, the problem of generalization does not exist as there is no initial stage of overgeneralization. A linking rules learner is predicted to always arrive at the adult grammar once they have learned the relevant conditions on each semantic verb class. It is worth noting here that previous work (e.g., Bowerman 1982) emphasizes the robustness with which children productively make causative errors, and this fact is not accounted for under an approach where the learner is constrained from the beginning.

In the description of the linking rules account described above, we see that Pinker 1984, 1989 suffers from the consequences of innate mapping between the syntax and the semantics. An innate linking rules account does not predict the causative errors productively produced by children. By eliminating the innate linking rules aspect using the Active Mapping Model of Language Learning proposed in Chapter 1, the child errors can be readily accounted for. The Active Mapping Model does not suffer from the same empirical problems as the linking rules account, even though the Active Mapping Model assumes less than an innate mapping model.

Additionally, in the same vein as Pinker 1989, Ambridge et al. 2008, 2011 also argue that children use the semantics of a verb to retreat from overgeneralization. As opposed to the linking rules approach in Pinker 1989, Ambridge et al. 2008 argues that retreat from overgeneralization is obtained through the learner statistically determining whether a verb can occur in a given syntactic frame. In a series of experiments, they show that children are sensitive to finer grained semantic distinctions between verbs, and are more likely to apply the causative alternation to intransitive verbs that are of the right semantic type, e.g., verbs that require external causation.

Throughout the rest of this chapter, I show that the errors are a result of a generalizable rule at earlier stages of the acquisition process. Retreat from overgeneralization errors are due to the change in generalizability of the rule resulting from the learner's growing vocabulary. Before diving into the analysis offered in this chapter, I first discuss one other kind of approach proposed to capture the causative overgeneralization facts. This type of approach falls under the general umbrella of indirect negative evidence, which I discuss in detail in the following section.

3.3.2. Usage-Based Approaches

In this section, I describe usage-based approaches that use indirect negative evidence in order to account for the retreat from overgeneralization (Brooks and Tomasello 1999; Bowerman and Croft 2008). Unlike models that assume linking rules between the verb semantics and verb argument structure, these approaches rely on the input data to a higher degree. Some usage-based models (e.g., Bowerman and Croft 2008) acknowledge that children's causative errors are productive. The causative errors are then resolved over time when a more frequent form of the verb in the input replaces the child's rules that stem from overgeneralization. This model relies on one crucial aspect, frequency. For verbs that are not overridden by a more frequent form, the learner may assume over time that the verb can only take the structure in which it occurs; i.e., over time, the learner assumes that if a verb has not occurred as a transitive, the transitive form of the verb is ungrammatical altogether. I elaborate on the role of frequency in these models below.

Usage-based indirect negative evidence approaches typically follow a statistical preemption (Bates and MacWhinney 1987; Bowerman and Croft 2008; Goldberg 1995; Pinker 1989) or entrenchment (Ambridge et al. 2008; Braine and Brooks 1995) model. I discuss statistical preemption approaches below, before describing entrenchment models of learning. Under this approach, there are two types of verbs that are overgeneralized in learning causatives: verbs that have a phonologically distinct, i.e., suppletive lexical causative form,

(e.g., *kill* and *die*), and verbs that do not (e.g., *disappear*). For suppletive verbs, the learner makes causative errors because the suppletive causative form of the verb has not yet been learned. For verbs that do not have a suppletive causative form, the make-causative is said to preempt the child's overarching hypothesis; for example, *make X disappear* preempts the overgeneralized form of *disappear X*. Statistical preemption falls under an indirect negative evidence approach because the learner must first assume that the transitive causative form of a verb is not possible in order to preempt it with another frequent grammatical form in the input.

Another line of work proposes the entrenchment of verb argument structure (Alishahi and Stevenson 2008; Ambridge 2013; Ambridge and Lieven 2011; Ambridge, Pine, Rowland, Jones, and Clark 2009; Ambridge, Pine, Rowland, and Young 2008). For instance, Ambridge et al. (2008, 2009) argue that the possible argument structures for a verb are stored each time they occur, and the more times the verb occurs, the higher the probability that the verb will only be used in the argument structure frames in which it has been encountered. Thus, when the child hears a verb like *fall* occur only as an intransitive, over time, the probability of the verb occurring as a transitive decreases until the learner assumes it can only occur as an intransitive. Furthermore, verb semantics also play a role in addition to entrenchment under this analysis, and the process of entrenchment is also claimed to include the entrenchment of semantic classes under some proposals (e.g., Ambridge 2013). The idea behind using entrenchment as a learning mechanism for the causatives problem is that children are said to learn the semantics of verbs through the linguistic environments they are found in. In other words, the verbs are entrenched in their meaning by their semantic environment. This semantic entrenchment constrains the learner from the beginning, and the learner assumes that the other verb uses in meaning are impossible. The notion of entrenchment, hence, involves the use of indirect negative evidence, as the absence of one verb form is taken as evidence for its ungrammaticality.

As stated earlier, the frequency of a verb plays a key role under an entrenchment

approach. Ambridge et al. 2008, 2009, 2011 argue that intransitive only verbs that are more frequent in the input are rated as less acceptable when used as a transitive causative than less frequent verbs in the same class. For instance, in testing adults, Ambridge et al. 2011 find a difference in the acceptability between *fall* and *tumble* as causatives. Participants were found to rate the more frequent verb *fall* a 1.46 on a 7 point Likert scale, and the less frequent verb *tumble* a 2.68 out of 7. I note here that the aforementioned results only suggest a possible effect of verb frequency, and not a robust distinction in the structural properties of these verbs. However, in order to confirm that adults at least show effects of entrenchment, one needs to rule out whether these effects could stem from other semantic properties of the verb. I address this question in Section 3.3.3.

Furthermore, Ambridge et al. 2008 argue that other situational factors play a role; 5 and 6 year old participants rate the causative form of the novel verb meaning *laugh* significantly lower than its intransitive form ($p = 0.001$); however, they do not rate the causative forms of novel verbs of disappearing and falling differently from their intransitive forms ($p = 0.1$ and $p = 0.9$ respectively). 9 and 10 year old children always significantly preferred the grammatical forms over the ungrammatical forms. Ambridge et al. 2008 argue that these results indicate that the level of direct external causation determines the level of acceptability of causatives. In Section 3.4, I show how these facts follow from an account that generalizes from positive evidence.

Although this approach discusses the mechanisms through which children can retreat from positing an incorrect causative form to verbs that do not allow them, the analysis does not discuss the reason for this overgeneralization. Why do children propose a rule for verbs that do not have a causative form in the first place? A linking rules approach addresses this problem by claiming that younger children have not fully mastered the specific properties of each semantic class, but under an entrenchment approach, the reason for overgeneralization is not entirely clear.

Now that we have discussed how indirect negative evidence has been proposed to ac-

count for the retreat from overgeneralization, we can evaluate whether these approaches can account for the problem at hand. In the following sections, I provide arguments against the use of indirect negative evidence to account for this problem by showing that these approaches do not address the productivity of the causative errors, nor are they able to adequately account for retreat from overgeneralization. I also illustrate, using child data, that even if children could learn from indirect negative evidence in theory, their developmental trajectory suggests otherwise. In Section 3.4, I argue instead that children overgeneralize when there is sufficient positive evidence to form a productive rule that can apply to verbs for which they have not heard the causative form.

3.3.3. Evaluating Preemption and Entrenchment

In this section, I evaluate the two indirect negative evidence approaches, statistical preemption and entrenchment, and show that these approaches proposed as a mechanism to retreat from the overgeneralization of causatives is not tenable. First, I discuss the possibility that children are using statistical preemption to learn verb argument structure below.

One way in which statistical preemption is formulated is under a semantic-based approach. Recall that Pinker 1989 argues that children initially erroneously learn verb semantics, and as a result, make causative errors. For instance, when children produce **go me to the bathroom*, they mean to say *take me to the bathroom*. Therefore, when children learn *take*, they will replace the erroneous form of *go*, and the causative errors will cease to occur. This kind of a statistical preemption approach also predicts that verbs that are more frequent should be learned sooner than more infrequent ones. More frequent grammatical forms can preempt infrequent ungrammatical forms. Moreover, a verb needs to be heard many times in the input in order for the learner to determine which forms of the verb are so infrequent that they can be considered impossible in the language. Since verbs that occur more frequently are expected to be learned earlier, errors with relatively

frequent verbs are predicted to cease sooner than less frequent verbs. Even from a perfunctory glance, this expected difference between verbs of varying frequency is not found. We have already seen that children produce causative errors with verbs as frequent as *go*, which is said to be used in place of *take*, another frequent verb. These verbs are more frequent than another possible suppletive pair *die* and *kill*, and yet, children produce these errors around the same time.³ Examining the child production data in CHILDES returns several examples of children producing these potentially suppletive pairs, *go* and *take* and *die* and *kill* at the same time. Moreover, children use *take* and *kill* correctly well before the age when they are known to make overgeneralization errors. Some early examples of *take* and *kill* in child speech are illustrated in (33).

- (33) a. I take it back (2;2, Shem, Clark corpus)
b. I kill monsters (2;6, Ross, MacWhinney corpus)
c. I'm gonna take you to the hospital (2;7, Peter, Bloom70 corpus)
d. I'm gonna take a gun and kill a fish for dinner (2;10, Abe, Kuczaj corpus)

As mentioned in the previous section, a crucial component of this model is the presence of a competing form of the same verb in the input that preempts the overgeneralized causative. For verbs that do not have a phonologically distinct lexical causative form in the input (like *stay*, but unlike *remember/remind*), the only other form in the input that can preempt causativized *stay* and *disappear* is the make-causative. However, the make-causative systematically carries a different meaning from the lexical causative (Ammon 1980; Bowerman and Croft 2008; Fodor 1970), and is its own productive construction that occurs independently. For example, Ammon 1980 discusses meaning contrasts between lexical and make-causatives, further supporting that these constructions are distinct. In the same vein, Shibatani 1973 discusses the meaning contrasts in some detail,

³Pinker 1989 argues that a preemption analysis under his account is not that of indirect negative evidence. The claim is that children simply never produce ungrammatical causatives of intransitive verbs because there is another causative verb with a similar meaning that can be used. In other words, the learner never truly loses the “ungrammatical” form. It is simply extremely infrequent in usage.

and divides them along the lines of the directness and indirectness in the meaning of the causatives. Similarly, Fodor 1970 also argues that make-causatives embed a syntactic structure whereas lexical causatives do not. Clark 1978 states that they are developed early and independently. Baron 1972 and Limber 1973 also find evidence for periphrastic causatives before the age of 3. Hence, Bowerman and Croft 2008 argues that the make-causative is a weaker cue in the input than a phonologically distinct lexical causative form. Errors with verbs like *disappear*, in that case, are predicted to persist longer than those with verbs like *die*, as there is no competing form to preempt it. However, Bowerman and Croft 2008 shows that this prediction is not borne out in examining the causative errors of two children C and E.

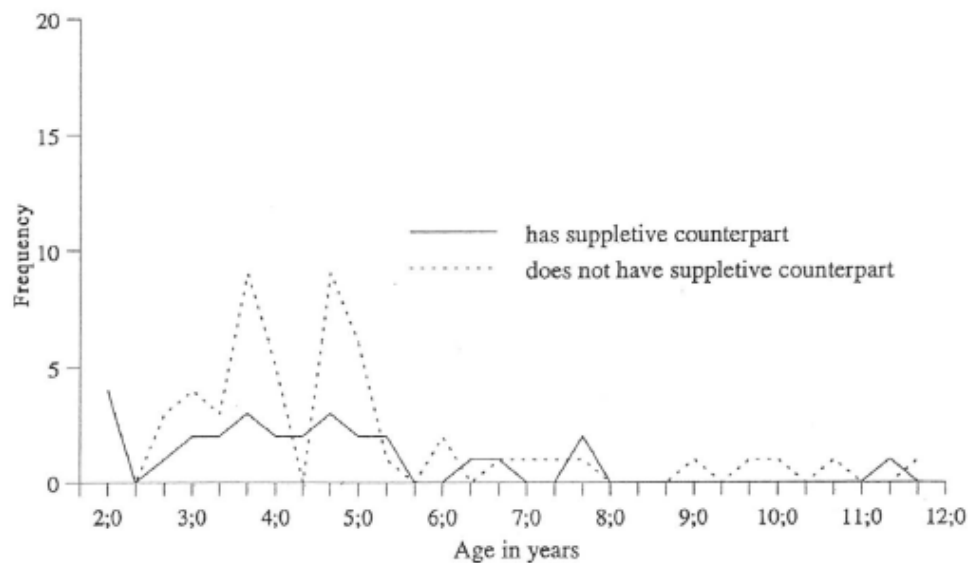


Figure 3.2: Child C's causative errors over time (from Bowerman and Croft 2008:297). No difference was found between the rates of causative errors with verbs that do not have a suppletive form and verbs that do.

As the graphs in Figures 3.2 and 3.3 show, there is no difference in the rate of declining errors between verbs that have a suppletive counterpart and verbs that do not for either child. Child C's errors for both kinds of verbs decrease at a similar rate, and for Child E, we find that there are in fact more errors with verbs that do have a suppletive

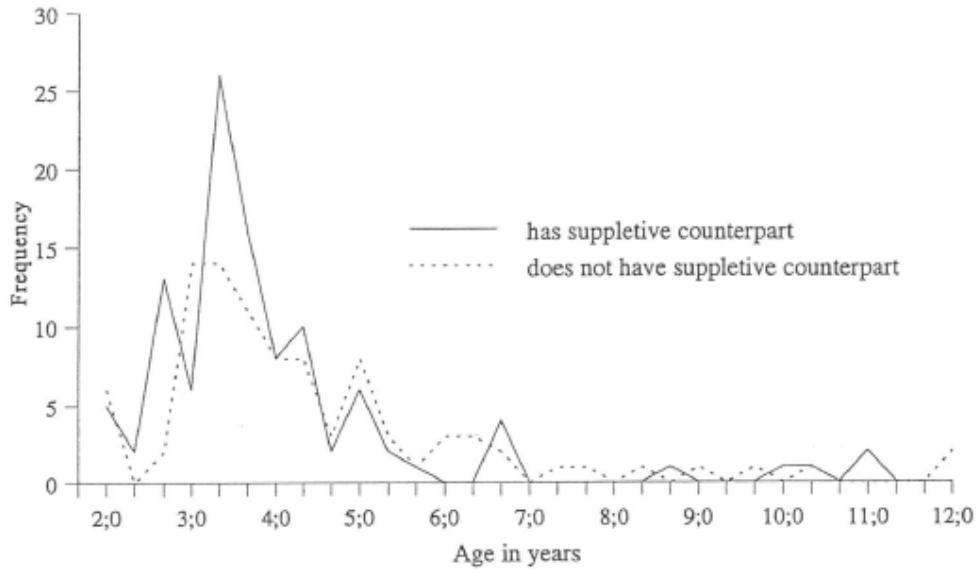


Figure 3.3: Child E's causative errors over time (from Bowerman and Croft 2008:297). No difference was found between the rates of causative errors with verbs that do not have a suppletive form and verbs that do.

counterpart between the ages of 3 and 4 years. These error rates in Figures 3.2 and 3.3 are our first indication that preemption using causative verb counterparts of the overgeneralized verbs may not be the right mechanism through which children retreat from their overgeneralization. If that were the case, then we would expect a difference between *disappear*-type verbs with no suppletive form and *die*-type verbs that have a suppletive form. Second, it is unclear whether make-causatives do preempt the causative forms of overgeneralized verbs at all. There is evidence from child data that children are aware of the make-causative early on, even for verbs that they overgeneralize. (34) provides some examples illustrating the productivity of make-causatives in child speech.

(34) Make-causatives in child data:

- a. someone made me go away (2;9, Ross, MacWhinney corpus)
- b. I'm going to make it disappear (3;2, Ross, MacWhinney corpus)
- c. Shoes may make you jump higher (3;9, Adam, Brown corpus)
- d. you stand up to make it go fast (3;10, Adam, Brown corpus)

- e. Mommy make it stay under dere (4;1, Adam, Brown corpus)
- f. I make my icecream disappear with no holes in it (4;3, Adam, Brown corpus)
- g. She's gonna make it die (5;0, Christy, Bowerman 1982)

As the shown in the examples above,, the verbs in (34) are exactly those verbs that children are known to make errors with. Moreover, the use of the make-causative overlaps with the time at which we notice the overgeneralization of the causative rule, indicating that children are aware of this alternative construction at the time when the errors are produced. Some examples of the causative errors made by Adam (Brown corpus) and Ross (MacWhinney corpus) are shown in (35).

- (35) Causative errors produced by Adam and Ross:
- a. and my mommy might break this and fall this (3;0, Ross, MacWhinney corpus)
 - b. I want to disappear it (3;3, Ross, MacWhinney corpus)
 - c. gon(na) fall him to pieces (3;7, Adam, Brown corpus)
 - d. how did it disappear this air out_of here (4;2, Ross, MacWhinney corpus)
 - e. how to go it ? (4;7, Adam, Brown corpus)

(35) shows that both Adam and Ross know the make-causative construction when they produce the lexical causative errors. The make-causative examples produced with verbs like *go* and *disappear* also illustrate that both children are aware of an alternative construction in place of a potential lexical causative form of these verbs. However, the make-causative does not preempt the lexical causative in these cases. Moreover, both of these forms for a single verb are clearly possible, and frequent in use. (36) illustrates some verbs that allow both the make-causative and the lexical causative forms, indicating that adults have a grammar in which both the make-causative and the lexical causative co-exist for the same verb.

- (36) Verbs that allow both lexical causative and make-causative forms:

- a. I made the ice melt / I melted the ice.
- b. I made the ship sink / I sank the ship.
- c. John made the house burn down / John burnt down the house.
- d. John made plants grow / John grew plants.
- e. Rita made the paper rip / Rita ripped the paper.

The examples above show that the make-causative and the lexical causative coexist. Therefore, it is not obvious why the make-causative would then preempt the hypothesized causative form of verbs. Since children are aware of the make-causative forms and use them frequently, as shown in (34) and (35), a preemption account cannot claim that children stop making causative errors by learning the make-causative for those verbs later on. In sum, our findings from child production data and the findings of much previous research indicates that a statistical preemption analysis for the acquisition of causatives is not tenable.

The alternative to preemption under indirect negative evidence approaches is entrenchment. Brooks and Tomasello 1999 argues, for instance, that children are more likely to use an intransitive frame for novel verbs heard in an intransitive frame; i.e., they are more likely to assume that a verb is confined to the syntactic constructions it occurs in. Entrenchment of a form then predicts that a verb heard in an intransitive frame, such as *die* or *disappear*, are unlikely to be used as transitives in the first place. This constraint on the verb's argument structure also has the effect of deeming child errors insignificant because they do not result from a productive generalization. As we have seen in the child data, these errors are produced with a wide variety of verbs. Children are visibly aware of a pattern in their language, which they are exploiting. In dismissing this pattern, we would be overlooking the possibility that children are principally forming productive rules in learning the grammar of their language.

In addition to the aspect of conservative learning implied by the entrenchment account, the prediction made by this approach regarding the kinds of errors children produce

is not borne out. An entrenchment approach, which claims that children cease to make causative errors once they've heard a verb used in an intransitive frame many times in the input, predicts a difference in the production of errors between verbs of varying frequencies. The child is expected to retreat from using an intransitive verb in a causative frame earlier if the verb occurs with a relatively high frequency in the input. This prediction can be tested by examining the causative errors made by the children in CHILDES, and their frequency in child-directed speech. Ross from the MacWhinney corpus (MacWhinney 2000), for instance, makes several causative errors. Some of the causative errors that Ross produces are shown below:

- (37) a. maybe I can fall it down the stairs (3;3)
b. and are you going to stay me at my new school at Pittsburgh (3;5)
c. to go it down my tummy (3;11)

As the examples above illustrate, Ross makes causative errors with different kinds of intransitive verbs. These verbs vary in terms of their frequency. To test whether these verbs were indeed not heard as an intransitive enough times in the input, I conducted a corpus search of these verbs in child-directed speech. To do so, I examined the input data of each caregiver in the corpora available for North American English. This resulted in a total of about 6 million words of speech. Out of the 6 million words of combined data, *disappear* occurs 153 times overall. *Stay* occurs a total of 2,662 times, *fall* a total of 2,819 times, and *go* occurs 55,689 in total. The results of this corpus search are summarized in Table 3.2. However, in spite of the differences in their frequency, the errors we find for Ross all cluster together around the ages of 3-4, and there is no evidence that the child stops producing errors from these verbs purely on the basis of hearing them intransitively a number of times. If the entrenchment hypothesis were true, we would not expect children to be making errors with verbs like *go* that occur over fifty thousand times in the input data. If it was true that a verb needed to be heard that many times in order to learn its argument structure, we would expect children to be making causative errors with verbs

like *disappear*, which occur a fraction of the time, well until their adult years. As we are aware, this is not the case.

Table 3.2: Raw frequency of pure intransitive verbs in child-directed speech.

Verb	Raw frequency in CHILDES
<i>disappear</i>	153
<i>stay</i>	2,662
<i>fall</i>	2,819
<i>go</i>	55,689

In the next section, I discuss an alternative learning mechanism to both statistical preemption and entrenchment. I argue that children make generalizations of the properties of verbs from sufficient positive evidence in the input, and they retreat from these generalizations when there ceases to be sufficient positive evidence in the input.

3.4. Learning Causativization

In this section, I describe how the learner uses conceptual and structural cues in their language in order to arrive at the adult grammar. The literature on child verb learning argues that children make use of both structural information and conceptual cues in verb learning (Gleitman and Gleitman 1992; Gleitman and Landau 1994; Naigles and Kako 1993; Pinker 1984, 1987), and the two are often linked to one another. These verb learning approaches are argued to be relevant for learning causative verbs as well (Ambridge, Pine, Rowland, and Young 2008; Naigles 1990, 1996; Naigles and Hoff-Ginsberg 1995; Pinker 1989). Here, I also describe how structural and conceptual cues aide the child in acquiring the different kinds of intransitive verbs. We delve into how the learner uses notions such as intentionality to distinguish between unaccusative and unergative verbs, as this distinction is relevant to the acquisition of causatives. Once the learner has identified the

two types of intransitives, I show how they generalize the causative rule from sufficient positive evidence available to them in the input. Sufficient evidence here is determined by the Sufficiency Principle proposed by Yang 2016, as previously described in Chapter 1. I also demonstrate how, as the learner acquires more verbs, there ceases to be sufficient evidence for the causative alternation rule given their increased vocabulary size, and the learner no longer generalizes. Furthermore, I argue that the motivation behind subdividing verbs into smaller classes may arise from the need to form productive rules in language (cf. Yang 2016); i.e., the learner searches for subclasses of verbs to identify a class where the causative alternation rule is in fact generalizable.

The learning model that we use to describe the acquisition of causatives was previously discussed and presented in Chapter 1. This Active Mapping Model is repeated again in Figure 3.4 below. Recall that in the Active Mapping Model, structural and conceptual cues are independent, and the two parts of the model are not inherently linked to one another.

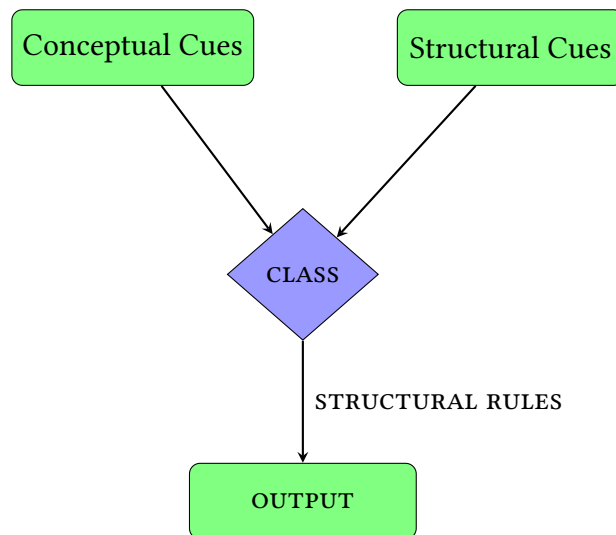


Figure 3.4: The Active Mapping Model of Language Acquisition, and a summary of how verb classes are acquired. Conceptual and structural cues both play a role in acquiring verb classes, but only structural cues can apply at the class level.

3.4.1. Learning Verb Classes from Structural and Conceptual Cues

This section describes the concepts and structural information used by the learner in acquiring verb classes. I first discuss how unaccusative and unergative verbs are differentiated, as this distinction plays a role in the overgeneralization of the causative rule. The causative alternation is a property that is typically observed for unaccusative verbs (e.g., Levin 1993). The adult grammar shows a productive causative rule for some subclasses of unaccusatives, such as the manner of motion class consisting of verbs like *roll* (see also Brooks and Tomasello 1999). However, not all kinds of verbs show this rule productively. For example, verbs of inherently directed motion do not generally display this property (e.g., *rise*, *come*). In showing how the learner acquires these subclasses where the causative alternation rule is productive, I address the conceptual and structural cues that they are attuned to. I primarily focus on intentionality, causality, and motion in this section.

Languages are known to exhibit two kinds of intransitive verbs (Perlmutter and Postal 1984): unergatives and unaccusatives. Unergative verbs introduce an external argument, and optionally allow certain kinds of objects, while unaccusative verbs do not introduce an external argument. On the surface, these two kinds of verbs appear to both have a surface subject and one argument. The structural difference between them is that the surface subject of an unaccusative verb is an underlying object. Given these two types of intransitives, the learner is faced with the task of distinguishing between the two classes of verbs, as a host of properties are associated with each type. How children acquire this distinction in English is also a particularly interesting question, as there are limited or no overt structural cues that differentiate them. Learning the two classes of verbs is also relevant to the acquisition of causatives as it is primarily the class of unaccusative verbs that show the property of the causative alternation; unergative verbs generally do not display this property. Consequently, the generalization of the causative rule that children make cannot stem from considering the entire class of unergative verbs as well. In order

for the rule to become a candidate for generalizing, the learner must have identified the distinction between the two types of verbs. Below, I describe how children make this differentiation between unaccusative and unergative verbs.

I argue that children use intentionality as a differentiating factor between unaccusative and unergative verbs. Unergative verbs in their intransitive frame are known to typically show intentionality (e.g., *Jamie ate*), whereas unaccusative verbs in their intransitive frame lack intentionality (e.g., *Jamie fell*). As stated in Chapter 1, there is some evidence in the literature that children are attuned to notions like intentionality early on (e.g., Tomasello and Barton 1994; Woodward 1998). For instance, Woodward 1998 uses a visual habituation paradigm to show that children distinguish between animate and inanimate agents. In these studies, infants between the ages of 4 to 11 months were habituated to the motion of a hand grasping or a rod touching one of the two objects placed in front of them. In the testing phase, the infants were shown the hand or rod grasping or touching a different object. The results showed that infants in the hand condition, but not the rod condition, looked longer at action when there was a new goal with an old path versus in trials that showed a new path with an old goal. These results indicate that infants as young as a few months old are able to encode the goal-oriented nature of actions. As intentionality as a concept is known to learners, I use intentionality here as a way of singling out unergative verbs from unaccusatives. Intentionality also comes into play in the following chapter, where I discuss the acquisition of passives. For now, I use it to distinguish between the two types of intransitive verbs.

Before moving onto the analysis, let us first discuss other cues that may be available in learning unaccusative and unergative verbs. It has also been argued that telicity sets apart unaccusative verbs from unergatives (Borer 1994; Grimshaw 1990; Levin and Hovav 1992; Tenny 1994; Tenny 1987; Van Hout 1996; i.a.), and moreover, that children use structural properties like auxiliary selection to distinguish between the two kinds of intransitive verbs (Van Hout 2013). In examining how children acquire unaccusative verbs, Van Hout

1993, 1996, 2004, 2013 propose that Dutch-learning children differentiate unaccusative verbs from unergatives along the lines of telicity. In this work, auxiliary selection is argued to lead the learner to acquire the host of properties that generally fall out from this distinction. Crucially, Dutch, the language investigated in this work, does show structural properties that lead the learner to identify the distinction between the class of intransitive verbs.

This key cue in Dutch, auxiliary selection, is unavailable in English. In Dutch, unaccusative verbs occur with *zijn* 'to be' in the perfective, whereas unergative verbs occur with *hebben* 'to have'. This fact is precisely what makes the question of the acquisition of the two kinds of intransitive verbs interesting in English. However, the syntactic representations of unaccusative and unergative verbs in English are still distinct. This distinction is observable from some of the standard diagnostics of unaccusativity. For instance, unaccusative verbs in English cannot be pseudo-passivized. Some examples of unaccusative versus unergative verbs are provided in (38) and (39). Additionally, let us also consider the following pairs of sentences from Perlmutter and Postal 1984:

- | | | | |
|------|----|-------------------------------------|----------------|
| (38) | a. | The gorilla sat on the desk. | [Unergative] |
| | b. | The desk was sat on by the gorilla. | [Unergative] |
| (39) | a. | The lamp sat on the desk. | [Unaccusative] |
| | b. | *The desk was sat on by the lamp. | [Unaccusative] |
- (Perlmutter and Postal 1984:103)

The examples in (38) and (39) all use the verb *sit*, but we get different results from passivization when the subject is *the gorilla* or *the lamp*. The only difference between the two pairs of sentences is that when the subject is *the gorilla*, the sitting action is intentional. When the subject is *the lamp*, the sitting action lacks intentionality. These pairs of sentences correspond to the unergative and unaccusative use of the same verb *sit*. Thus, in learning the two kinds of intransitives in English, intentionality plays a role.

It should be noted here that I use intentionality instead of agentivity as a conceptual cue for the learner, although the two are closely related. I use intentionality instead of agentivity because agents at times have been argued to greatly overlap with animate subjects in the language acquisition literature (e.g., Becker 2006). However, although there is some correlation between agents and animacy, animacy is not a reliable indicator of unaccusativity to the learner. Consider the following unaccusative verbs with animate subjects.

(40) John fell.

(41) The boy vanished.

(42) The rabbit disappeared.

The sentences provided in (40)–(42) are all unintentional intransitives, and the surface subject is non-agentive in these examples. However, the subjects of these sentences are animate, even if they are non-agentive. The examples above indicate that children cannot derive agentivity reliably from animacy, although there is a strong correlation between animacy and agentivity. Therefore, I propose that children use intentionality to distinguish between unergative and unaccusative verbs.

Now that we have discussed how children learn the larger set of unaccusative verbs, we can discuss how children learn finer distinctions within this large class. We recall from earlier sections that children and adults both productively use the causative rule when there is direct external causation (e.g., Maratsos et al. 1987), and moreover, that manner of motion verbs show this property as well (e.g., Brooks and Tomasello 1999). Here, I present findings that argue that children are attuned to notions such as causality.

It has been shown that children can identify the causality component of a verb from multiple syntactic frames (Landau and Gleitman 1985; Naigles 1996; Naigles and Hoff-Ginsberg 1995). Naigles (1996) shows that learners cue into the meaning of a verb, not only from causative frames, but also from other syntactic frames that may denote a causative

meaning. It has also been argued in this line of work that children are able to differentiate between contact verbs like *touch* and verbs that reflect causation in the world like *break*, suggesting that the notion of causality is truly present. These findings are relevant to the work presented here, as the learner arrives at a grammar that singles out verbs that indicate direct external causation.

Structural information also comes into play in learning causative verbs. For some abstract verbs, if a causative form exists in the language, it must be learned from the linguistic input. This fact is clear from cross-linguistic differences. A verb such as *arrossire* ‘blush’ has a direct causative form in Italian (e.g., Levin and Hovav 1992), but not in English where sentences like **the sun blushed my cheeks* are ungrammatical. Similarly, the Hindi verb *khilaanaa* ‘to make bloom’ or ‘to make blossom’ has a lexical causative form, which does not exist in English. This verb in English is said to be internally caused, and therefore, does not undergo external causation (e.g., Pinker 1989), but this claim does not hold for this verb in Hindi. The examples from Hindi and Italian indicate that the same event can be encoded as externally or internally caused depending on the language.

However, as discussed earlier, the possibility of direct causation need not be learned from the input for all verbs, especially not verbs that involve physical contact such as *fold*. In these cases, there may be situational evidence for a verb to be learned as involving direct causation, but we cannot do away with the structural information that supplements the learner’s knowledge for more abstract verbs denoting a direct causative relationship. In languages where there are structural and situational cues available to the learner, it has been shown that they are indeed useful in acquiring the grammar of their language (van Hout 2004). Thus, both situational and linguistic cues are indispensable to the learner in acquiring the finer meanings of unaccusative verbs.

In sum, studies on children’s knowledge of linguistic and situational cues indicate that children are attuned to concepts such as intentionality and causality as early as 12 months (Carpenter, Akhtar, and Tomasello 1998; Woodward 1998; Woodward, Sommerville, and

Guajardo 2001). Leaving aside the question of whether children are innately endowed with the ability to identify cues such as causality and intentionality, we can at least use the findings that children are aware of these concepts early on. Now that we have seen which structural and semantic cues are relevant to the learner in acquiring the semantic and syntactic properties of verbs, we can move on to the question of how children overgeneralize the causative alternation rule from positive evidence in the input, and how they eventually retreat from this overgeneralization.

3.4.2. Overgeneralizing The Causative Alternation

In this section, the prediction of whether children have sufficient evidence to overgeneralize the causative alternation rule is tested using corpus data. As stated in Chapter 1, I adopt the Sufficiency Principle (Yang 2016) here to determine when there is sufficient positive evidence in the input. The Sufficiency Principle states that in order for a rule to be generalizable, the number of exceptions to a rule for a class of N members must not exceed $N/\ln(N)$. Following the Sufficiency Principle, the causative errors only arise when the number of exceptions in the input are below the threshold to generalize this rule to other members of a given class. This hypothesis can be tested by investigating the causative errors made by children in corpus data (CHILDES, MacWhinney 2000). I examined the child production data of Adam (Brown 1973) and Ross (MacWhinney 2000) as two case studies of children who go through a stage of overgeneralizing the causative alternation. These two children are good candidates to test the claims made in this chapter as there is more data available for them than for most individual children in CHILDES, and both of these children make causative errors.

3.4.2.1. Test Case 1: Adam

First, I look at Adam's production data as a proxy for his vocabulary size. I examine each verb produced by Adam up to the point where he makes his first causative error, which

is at age 3;2. Then, I investigated whether there was evidence for a causative form in the input for the verbs produced by Adam. Each verb produced by Adam was checked in the combined North-American CHILDES input data to estimate the input he has likely received. Using this method, I determine whether the causative rule is productive for Adam given his vocabulary size.

Adam produces several causative errors starting at age 3;2. The errors noted for Adam often involve unaccusative verbs such as *fall* and *go*, among others. Some of the errors produced by Adam are shown in (43) below.

- (43) a. don't fall my head (3;2)
b. gon(na) fall him to pieces (3;7)
c. how to go it? (4;7)
d. if you can go it fast the pictures might run (4;7)

The causative errors for Adam extend until age 4;7; however, since the data for Adam ends at age 5;2, we cannot say for sure when these errors stopped occurring. The errors we do find indicate that the stage of overgeneralization lasts for as long as Adam's corpus.

To establish whether the causative errors arose in the child production data from sufficient positive evidence in the input, I examined Adam's production data for each verb used until the first unambiguous error was produced at age 3;2. Up to age 3;2, Adam produces a total of 208 verbs in CHILDES.⁴ Out of these 208 verbs, 110 of them occur as intransitive. From these verbs, I used intentionality as a guideline to distinguish between the intransitive verbs produced by Adam. As seen early on in this chapter, intentionality not only serves as a baseline for the divide between unergative and unaccusative verbs, but it is also a cue that children recognize from infancy. Since intentionality is a cognitively salient conceptual cue, it can be used here in analyzing whether the learner deems the causative rule productive given their vocabulary and the verb classes the causative rule

⁴The complete list of Adam's verbs can be found in the appendix.

potentially applies to. As shown later on, it is indeed primarily the class of unaccusative verbs, or unintentional intransitive verbs, that undergo the causative alternation rule.

For Adam’s verbs, intentionality served as a rough divide for differentiating most unergative verbs from unaccusative verbs. After identifying verbs showing intentionality, 51 verbs that lacked intentionality in an intransitive frame remained. These verbs included two verbs, *laugh* and *cry*, which are generally classified as unergatives in the adult grammar (Perlmutter and Postal 1984).⁵ The remaining 59 intransitive verbs were intentional intransitives. Adam’s verbs and whether the causative alternation rule is productive for Adam within a verb class is summarized in Table 3.3.

Table 3.3: Total number of Adam’s verbs sub-divided according to the way in which the verbs were used in child-directed speech. The causative alternation rule is productive for Adam for the class of unintentional intransitives, but not for the class of intentional intransitives. The threshold of generalization was calculated using the Sufficiency Principle.

Verb Type	# of Verbs	Causative Verbs	Threshold	Productive?
transitive	98			
intentional intrans.	59	9	44	NO
unintentional intrans.	50	38	37	YES
Total	207	49		

The table presented above summarizes the findings from the corpus search. The causative alternation rule is productive for Adam for the class of unintentional intransitives, but not for the class of intentional intransitives. Below, I further discuss the details of how the rule comes to be productive for Adam.

Recall that Adam’s verbs were counted up to the first causative error by hand. The verbs were then divided according to whether they were used as intransitives or only transitives in the CHILDES input data. For this step, data from each caregiver in the

⁵These verbs can eventually be learned as unergatives, under some readings, *laugh*, for instance, is clearly intentional (e.g., *Susan laughed at his foolishness*). For now, these verbs are grouped together with the class of unintentional intransitives, as the learner might not categorize these verbs that way immediately.

North American English CHILDES corpora were examined, rather than just the child’s caregivers’ speech. The child’s production data within a single corpus can be used to determine whether the child knows a verb, but because words in the input follow a Zipfian distribution, they may not occur in every possible syntactic frame. The Zipfian distribution results in many verbs only occurring only a few times, or even just once in the input data. To ensure that the corpus size was large enough for the verbs to occur enough times, Adam’s verb’s argument structures were compared with the combined CHILDES input data. The result was 109 verbs that occurred as intransitives and 98 verbs that occurred as transitives. The intransitive verbs were then divided along the lines of intentionality. The way in which the verb classes were obtained is summarized in Figure 3.5. For each intransitive verb, I checked to see whether the verb occurred as a transitive causative in the input. As shown in Table 3.3, 38 out of the 50 unintentional intransitives occurred in a transitive causative frame, and only 9 out of 59 of the intentional intransitives occurred in a transitive causative frame.

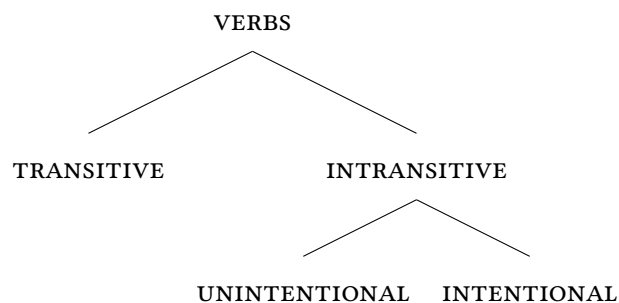


Figure 3.5: Classification of verbs using structural and conceptual cues.

Out of the unintransitive 50 verbs, 12 of them did not occur in a causative frame in the input data in CHILDES. Many of these verbs like *fall* do not have a causative form in the adult grammar. Following the Sufficiency Principle, a class with 50 members requires at least 37 verbs to show a property in order for the property to be generalizable. Here, 38 verbs show the causative alternation, and hence, the rule is generalizable to the other

members of the class, and Adam makes causative errors.

In total, Adam used 59 verbs that were classified as intentional intransitives. Out of these verbs, only 9 of them causativize in the input. Therefore, Adam does not overgeneralize the causative rule to intentional intransitives, as the Sufficiency Principle requires 44 verbs to follow the rule in order for it to be used productively. This result is consistent with what we find in Adam's production data. Adam makes no causative errors with unergative verbs. Moreover, as only a fraction of the intentional intransitive verbs have a causative form, it would be unexpected that the child would generalize the causative alternation rule if they were taking into account the entire set of intransitive verbs. Collapsing the two classes of intransitives into one results in only 47 out of 108 verbs that causativize. The number of causatives then falls below the threshold of the Sufficiency Principle. However, the total number of causative verbs falls also well below 50%. Given this data, most models would not make predictions for the learner to generalize when only a minority subset of a class follows a certain pattern.

3.4.2.2. Test Case 2: Ross

Another case study of the causative overgeneralization was obtained from Ross's data in the MacWhinney corpus. Like Adam, Ross also produced causative errors around the ages of 3-4 years. Some of these causative errors are illustrated in (44).

- (44)
- a. and my mommy might break this and fall this (3;0)
 - b. I want to disappear it (3;3)
 - c. and are you going to stay me at my new school at Pittsburgh (3;5)
 - d. to go it down my tummy (3;11)
 - e. how did it disappear this air out of here (4;2)

The last error found for Ross is around age 4;2. The errors here are clustered around the ages of 3 and 4. Similar to what was found for Adam, Ross makes causative errors with

unaccusative verbs.

To test whether there was sufficient evidence in the input in order for Ross to generalize the causative rule, I examined the verbs in Ross’s production data, and compared them with the CHILDES input data. I counted each verb that Ross produced by hand up to the point when he makes his first error, which was at age 3;0. For Ross, I obtained a total of 121 verbs that were produced up to that point.⁶

Ross’s verbs were also categorized based on transitivity. Out of 121, 66 verbs occur as plain intransitives in CHILDES. 42 of these intransitives showed intentionality. The remaining 24 verbs were intransitives lacking intentionality, which also included *cry*, an unergative verb. 55 of the 121 verbs were transitives. The different verb types are summarized in Table 3.4.

Table 3.4: Total number of Ross’s verbs sub-divided according to the way in which the verbs were used in child-directed speech. The causative alternation rule is productive for Ross for the class of unintentional intransitives, but not for the class of intentional intransitives. The threshold of generalizability was calculated using the Sufficiency Principle.

Verb Type	# of Verbs	Causative Verbs	Threshold	Productive?
transitive	55			
intentional intrans.	42	7	31	NO
unintentional intrans.	24	18	16	YES
Total	121	25		

As seen in Table 3.4, 18 of the 24 unintentional intransitive verbs occurred with a causative form in the input in CHILDES. A class with 24 verbs requires 16 for the rule to be generalizable, 18 verbs is above that threshold. Thus, we find that there is sufficient positive evidence for the causative alternation rule in the child-directed input for the verbs that Ross has learned up to this point. This sufficient evidence leads him to overgeneralize the causative rule to other verbs in the class. Consequently, Ross makes causative errors.

⁶The complete list of Ross’s verbs can be found in the appendix.

Additionally, we also see that there is insufficient evidence for the rule to be productive for the subclass of intentional intransitives. Only 7 out of the 42 verbs in this category occur as causatives in the child-directed input. Moreover, the two intransitive types cannot be collapsed into one. If the two intransitives types were treated as one group, for a total of 66 intransitive verbs, only 25 would have been witnessed in the transitive causative frame. Once again, this number falls below the threshold of the Sufficiency Principle. With less than 50% of the intransitive verbs undergoing the alternation, however, it would also be unexpected under most learning models for the learner to generalize the rule when it is not followed by the majority of the members of that class.

3.4.3. Retreat from Overgeneralization

Now that we have seen how overgeneralization occurs, we can discuss how children eventually retreat from their overarching hypothesis to arrive at the adult grammar. I claim here that the errors cease to occur when there is insufficient positive evidence in the input for the learner. To find evidence for the unproductivity of the rule in later stages of acquisition, we must estimate the growing vocabulary of the learner. To test whether the number of verbs that undergo the rule indeed decreases with increase in vocabulary size, I examined the properties of verbs listed in Levin 1993. I then found the frequencies for these verbs in CELEX, which is an adult corpus. Next, I looked for the verb with the lowest frequency in CELEX that occurred in CHILDES. This way, the number of verbs the child comes to know in later stages of acquisition can be estimated. The verb with the lowest frequency that occurred in CHILDES is *germinate*, which has a frequency of 1 per million. Hence, I only considered unintentional intransitive verbs from Levin 1993 that occur with a frequency of 1 per million or higher, which resulted in a total of 261 verbs. Out of the 261 verbs, 65 only occur as intransitives; i.e., only 196 of the verbs have a causative form. According to the Sufficiency Principle, a class with 261 members requires at least $N - N/\ln(N) = 214$ verbs to follow the rule. 196 is insufficient evidence for the learner to generalize,

and therefore, they cease to apply the rule to verbs for which they have not encountered the causative form. Notably, retreat from generalization is expected even if all of the 196 verbs are learned by the child. A more likely scenario is that the learner may not know all of the 196 verbs, which may result in the number of causative verbs falling much lower than the Sufficiency Principle threshold described here.

It is important to note here that learning the verbal properties cannot stop at the stage where the causative alternation rule is unproductive. We recall from earlier sections that there is evidence showing that both adults and children use novel verbs that indicate external causation in a causative frame (Maratsos et al. 1987). These results show that the grammar does have productive rules for causativization. The approach I have taken here allows us to address how children end up with a productive rule for some verb classes, but not for all intransitive verbs overall. Recall that the goal of the learner is to find rules and patterns within their language to facilitate mastery of the grammar. Once the learner reaches the stage where the rule for forming causatives is not generalizable, they are motivated to look for further subclasses where the rule might in fact be productive. Two subclasses that are known to allow the causative alternation productively are the manner of motion verb class (e.g., *the ball rolled/John rolled the ball*) and verbs that lack internal causation (e.g., *the house burned down/John burned the house down*). I claim that children then learn these subclasses when the rule is not generalizable to the larger set of unintentional intransitive verbs.⁷ The overall learning mechanism is illustrated in Figure 3.6.

As seen in Figure 3.6, the learner first posits a linguistic class that has certain properties associated with it. In this case, the learner posits a class of unintentional or intentional intransitives. The rule under consideration for this class is the causative alternation rule. The learner then computes, according to the Sufficiency Principle, whether or not the

⁷The same reasoning does not stop the learner from forming a productive subclass within the intentional intransitives. The intentional intransitive verbs that undergo the alternation are those like *walk* or *gallop* in uses like *I walked the dog* or *I galloped the horse*. These uses are specific in that they involve direct external causation on the object doing the action, and these verbs of motion may form a class of their own.

rule observed is a productive or unproductive one. If the rule is productive, the learner generalizes. If the rule is unproductive given the learner's vocabulary, the learner seeks subclasses where the rule may be productive.

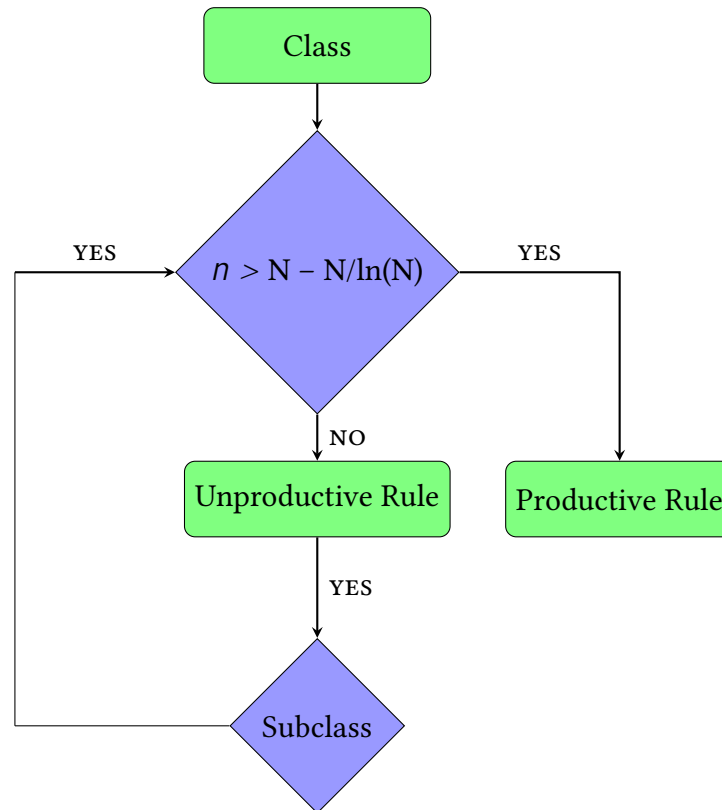


Figure 3.6: Summary of the learning mechanism employed by children. Rules are determined based on sufficient positive evidence in the input as determined by the Sufficiency Principle. If the rule is unproductive, the learner seeks subclasses where a productive rule may be found.

The results found for Adam are compatible with the account proposed above. Recall that Adam produced causative errors as long as the production data for him were available. He made errors with verbs like *fall* and *come*, which are motion verbs. Using motion verbs as a test case, we find that Adam in fact did not have the motivation of productivity to further divide his motion verbs into subclasses of directed motion and manner of motion, as the causative alternation rule was productive for him. Examining Adam's data reveals a total of 13 unintentional intransitive motion verbs: *go, come, fall, shake, move, turn, drop,*

pop, roll, slip, wind, swing, and rock. Out of these verbs, only 3 of them do not occur as a transitive causative. Therefore, even if we assume that Adam is able to form a subclass of motion verbs, the rule for the unintentional intransitive motion verbs is productive. However, we have no evidence as yet for why a subclass of motion verbs would be formed. Seeking productivity for a rule could serve as one kind of motivation driven by learning mechanisms. In the data examined, Adam does not have the motivation to propose a further subclass of manner of motion verbs where the rule may be productive. Even if Adam proposes a subclass of motion verbs, verbs that we classify as inherent motion may also be causativized as a result.

I have shown so far how the evidence available in the input influences the learner's decision to generalize a rule. There is one final point left to consider. Under some learning models, there is an innate mapping relationship between the structural and semantic properties of the verb. If the mapping between the semantics and the verb argument structure was innate, some of the child behavior observed would not be expected. If the mapping was innate, no child overgeneralization errors are predicted. However, child production data indicates errors with verbs like *fall* and *disappear*. The argument under such models is that the verb semantics are incorrectly learned by the child in that case. This explanation is ad-hoc, and does not account for the productivity of these errors.

Instead, I argue that the learner forms subclasses of verbs based on conceptual and structural cues, and any mapping between the syntax and the semantics is gained from the learner's linguistic experience. For each subclass, the learner then determines whether the causative alternation rule is productive. The details of the analysis proposed in this chapter are summarized in Figure 3.7.

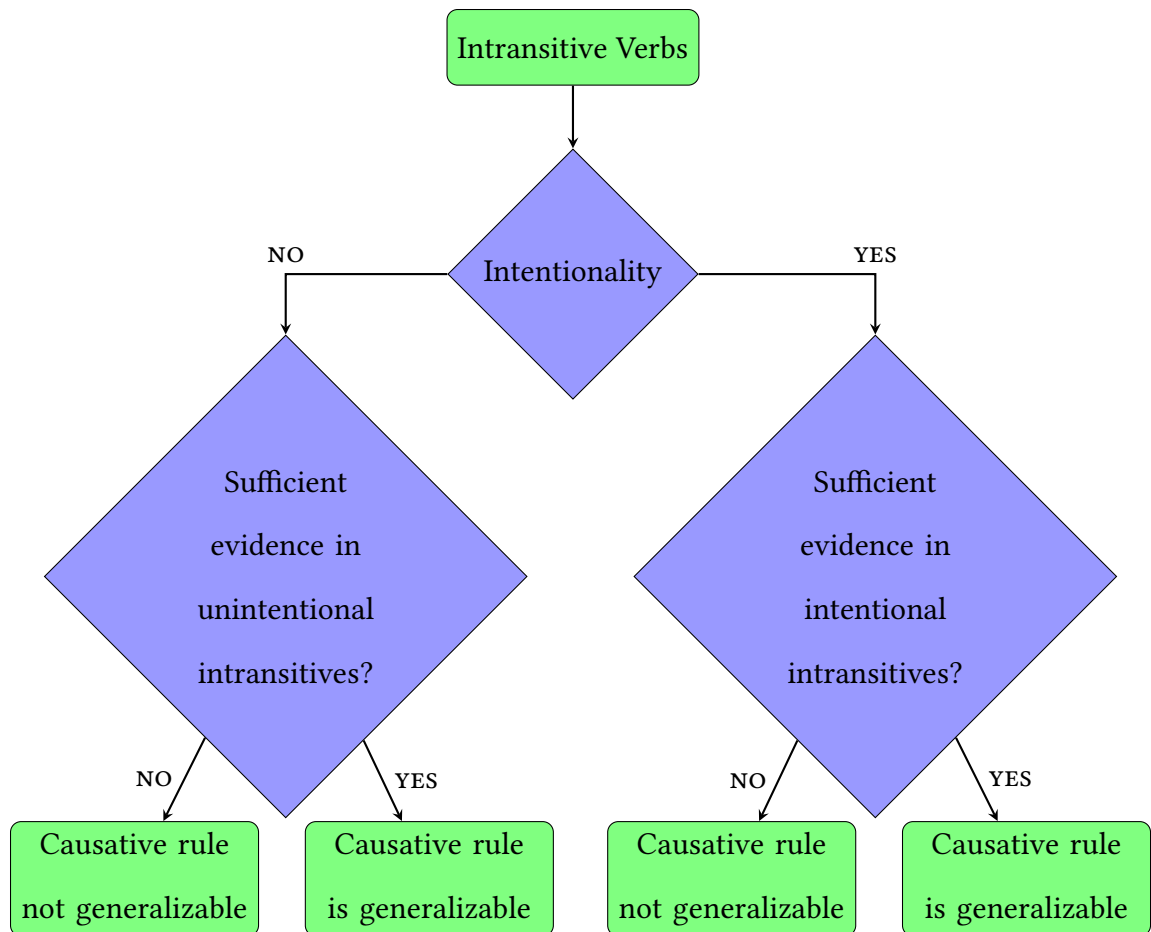


Figure 3.7: Summary of how the causativization rule is learned. Children first determine whether a given intransitive verb shows intentionality. Then, for each of those verbs, it is determined whether there is sufficient positive evidence to generalize the rule following the Sufficiency Principle. The causative alternation rule is generalizable when there is sufficient evidence, and it is not generalizable when there is insufficient evidence.

3.5. Interim Summary

In this chapter, I have provided a theory of how the child can learn the causativization rule in their language from positive evidence in the input. I have also illustrated how both conceptual and structural cues are used by the learner to acquire verb classes in their language. I have also argued that the input data provides the child with sufficient evidence to postulate a productive causative rule in their grammar. Retreat from overgeneralization occurs when there is insufficient evidence.

Crucially, it is not merely possible for us to assume learning from positive evidence, but absolutely necessary. I have shown here that evidence for indirect negative evidence in the input is sparse, and moreover, the learner does not appear to use indirect negative evidence reliably in acquiring verb argument structure properties in their language.

In this chapter, we also discussed whether the learner can acquire the causative alternation rule purely on the basis of verb semantics (Pinker 1989). It is problematic to assume that the semantics lead to a particular verb structure, as the relationship between the transitive causative and the verb semantics vary crosslinguistically (e.g., Pye, Loeb, and Pao 1996). For instance, the verb *fall* in English does not have a transitive causative counterpart, but the verb *girnaa* ‘fall’ does in Hindi where *giraanaa* ‘make fall’ is part of the grammar.⁸ Moreover, under this account, the eventual retreat from overgeneralization is either unclear or the result of indirect negative evidence.

The analysis presented in this chapter accounts for the facts about the acquisition of causatives that have been observed in the literature; however, one question that could be raised in the context of the analysis presented here is whether gradient judgments can be accounted for in this system. For instance, Ambridge et al. 2008 finds that adult participants are less likely to causativize a verb with *laugh*-type semantics than a verb with *disappear*-type semantics. Crucially, *laugh* in its intransitive form in Ambridge et al. 2008 is treated as an inchoative on par with intransitive *disappear* (Ambridge et al. 2008:87, footnote 1). Under our analysis, the gradient differences in grammaticality judgments fall out from verb classes formed using conceptual and structural cues.

Throughout the chapter, and indeed, throughout this dissertation, I have employed the Sufficiency Principle to determine whether a rule is productive or unproductive. I have used the Sufficiency Principle instead of assuming that a majority of the verbs following a rule is sufficient for the learner to assume generalizability. In the following section, I experimentally test whether the learner follows the Sufficiency Principle in generalizing

⁸Hindi *giraanaa* ‘make fall’ is a direct lexical causative, as indicated by the direct causative *-aa* morphology. Indirect causatives are marked with *-vaa*, as in *girvaanaa* ‘make/had fall’.

the causative alternation rule. The causative alternation rule is an ideal construction to test, as children overgeneralize the causative rule in natural speech.

3.6. Testing the Sufficiency Principle

Throughout the dissertation, and in this chapter, I have used the Sufficiency Principle (Yang 2016) to determine whether or not the learner has sufficient evidence to generalize a rule from positive evidence in the input. I also argue that overgeneralization errors stem from the generalizability of a rule. Children make errors when they use a rule productively, and apply it to forms that are exceptions to that rule. Eventually, the rule is not generalizable, and the learner retreats from this overgeneralization when the number of exceptions to the rule exceed the threshold determined by the Sufficiency Principle. As the Sufficiency Principle features centrally in this dissertation, I experimentally test the predictions it makes in this section. Here, I test whether children and adults follow the Sufficiency Principle when learning the rules of an artificial language.

Previous research on how children learn verb argument structure has argued for different approaches including statistical learning, syntactic bootstrapping, and semantic bootstrapping. In the following section, I briefly discuss experimental work that describes generalizing a rule through statistical learning mechanisms. I then present the experiments conducted in the current study that test the threshold of generalization. In Sections 5.3 – 5.5, I present results from three experiments, two on adults and one on child participants. Each experiment aims at capturing the point at which the participant overgeneralizes a rule from the input.

3.6.1. Background

In this section, I first discuss literature on studies conducted on how children generalize from the linguistic input more broadly. Work on regularization in child language acqui-

sition has found that children regularize forms in their own production even when there are inconsistencies in the primary linguistic input (Aslin and Newport 2012; Austin 2010; Hudson-Kam and Newport 2005, 2009). For instance, Schuler, Horowitz, Yang, and Newport 2017 find that children regularize inconsistent plural marking on nominals. When a particular nominal is used with a dominant marker between 40-67% of the time, but not 33% of the time, children regularize the dominant marker to produce it over 80% of the time in their own usage (Austin 2010; Schuler et al. 2017). This line of research suggests that regularization at the word level can occur even when the dominant marker only occurs 40% of the time. These results indicate that children make generalizations from positive evidence in the linguistic input. It should also be noted here for clarification that these studies do not have any bearing on the Sufficiency Principle. The Sufficiency Principle crucially refers to generalization across word types, and not regularization within a particular form, which is what the reported studies are testing.

Another line of work tests whether children are sensitive to entrenchment effects when learning verb argument structure (e.g., Ambridge 2013; Ambridge, Pine, Rowland, and Young 2008). When investigating the acquisition of causatives, Ambridge, Pine, Rowland, and Young 2008 and Ambridge, Pine, Rowland, Jones, and Clark 2009 argue that the absence of a causative form for a given verb serves as evidence against a causative form for the verb altogether. In other words, over time, if the learner does not hear a causative form for a verb, they assume that it does not exist. Under such a model of entrenchment, the higher the verb frequency, the more unexpected it is for the learner to overgeneralize a causative form for that verb given that the causative form for that verb has not occurred in the input. In their experiments, Ambridge et al. 2013 finds, for instance, that participants rate the ungrammatical causatives of less frequent verbs as more acceptable than the ungrammatical causatives of more frequent verbs. To address these claims, the experiments presented here also test the extent to which entrenchment effects are found in verb argument structure learning.

In addition to studying morphological marking in the nominal domain, artificial language studies have also been conducted to investigate how children learn verb argument structure (Aslin and Newport 2012; Braine et al. 1990; Gómez and Gerken 2000; Perfors, Tenenbaum, and Wonnacott 2010; Wonnacott, Newport, and Tanenhaus 2008). For instance, Wonnacott, Newport, and Tanenhaus 2008 and Perfors, Tenenbaum, and Wonnacott 2010, among others, investigate the acquisition of rules with respect to the causative alternation. These studies are directly relevant to the experiments described here, as these experiments also test the generalizability of the causative alternation using an artificial language paradigm. Wonnacott et al. 2008, for instance, conduct three experiments on adult participants testing whether participants would be able to acquire the causative verb argument structure given varying levels of positive evidence in the input. In their first experiment, they test 12 novel verbs: 4 that only occur as transitive causatives, 4 that occur only as intransitives, and 4 that alternate between the two. The experiment took place over 5 days, with one day devoted to vocabulary learning. They test participants on a production task and online comprehension task. A vocabulary test was also administered, in which all participants were reported to score 100%, indicating that the vocabulary of the language had been learned. The results of the first experiment show that participants had learned the correct argument structures of the verbs; i.e., participants did not overgeneralize the rule. Wonnacott et al. 2008 also conducts a second experiment testing 2 pure intransitive verbs, 2 pure transitive causative verbs, and 8 alternating verbs. This second experiment showed a stronger tendency for participants to overgeneralize, where the intransitive verbs were often used in the transitive causative form.

Under the analysis discussed throughout the dissertation, the difference between experiment one and two is unsurprising when the Sufficiency Principle is taken into account. By performing an ad-hoc analysis on the results obtained by Wannacott et al. 2008, we find that the number of alternating verbs did not meet the threshold of the Sufficiency Principle, as only 4 out of 12 showed the causative alternation. In contrast, in Experiment

2, 8 out of 12 verbs showed the causative alternation, which is right at the threshold of the Sufficiency Principle ($12 - 12/\ln(12) = 7.2$). The Sufficiency Principle can, thus, explain the tendency to overgeneralize that was found in the second experiment.

Finally, in a third experiment, Wannacott et al. 2008 test participants in two conditions: one where 7 out of 8 verbs occurred only as causative transitives and one verb only as an intransitive, and another condition where 8 verbs all occurred in both transitive and intransitive forms. In the first condition, participants learned the verb subcategorization accurately, whereas in the second condition, participants probability matched. The number of causative forms produced by the participants in the second condition were roughly the same number as what they heard in the exposure phase. These results are also unsurprising given the analysis argued for in this chapter. In their third experiment, participants were not exposed to a causative alternation rule; verbs simply took one form or the other. In the alternating condition, there is no room for overgeneralization as each verb tested took both structures. Now that previous research on generalization from the input has been discussed, the aims of the present study can be described in detail.

3.6.2. The Present Study

While Wonnacott et al. 2008 only test adults, the experiments presented in this chapter test both adults and children. The experiments described here may also make different predictions for the two groups, as the Sufficiency Principle might also only apply as a model of child language learning. This leaves open the question of how adults learn rules in their language, although there are some arguments that the Sufficiency Principle may apply to both adults and children (Yang and Montrul 2017). Some further, more minor differences between the present study and that of Wonnacott et al. is the structure of the artificial language. The present study and that of Wonnacott et al. differs in the word order used. While the agent was always included in their exposure phase, even for the intransitive verbs, the intransitive verbs in the present study only show a verb-patient or-

der. Moreover, the causative particle *-ka* in their study occurred at the end of the sentence, whereas in the current study, *-ka* attaches to the verb, similar to what we find in natural languages that show the causative alternation by adding the causative morpheme to the verb. Thus, the present study tries to mimic natural language learning more closely.

In the following sections, I present three experiments. Experiments 1 and 2 test adult subjects to investigate how they generalize rules, and whether the threshold for generalization is what is predicted by the Sufficiency Principle. The experiments also test whether adults overgeneralize by extending a rule to all members of a class that could potentially follow the rule. Experiment 3 tests the threshold of overgeneralization in child participants.

The experiments presented in this chapter directly test entrenchment effects as well. The Sufficiency Principle predicts that for a class with 10 members, at least 6 of the 10 must follow a rule in order for that rule to be generalizable ($N - N/\ln(N) = 5.7$). If 6 or more of the members show positive evidence of following the rule, then the learner may generalize the rule to the other 4. 5 or fewer members does not suffice. In order to test the Sufficiency Principle, I designed an experiment using an artificial language with 10 verbs in two language groups. One group consists of 5 out of the 10 verbs with a causative form, while the other showed 8 out of 10 verbs with a causative form. Participants exposed to the 8 out of 10 language are expected to overgeneralize, whereas participants in the 5 out of 10 language group should not.

A key component of the experiments presented here that equip us to test entrenchment effects is the variation in the frequency of the items used. The verbs in the experiment follow a Zipfian distribution, which means that there is a wide distribution in the frequency of the verbs. When a higher frequency verb occurs without a causative form, we can see whether participants treat it differently from a verb that lacks a causative form and occurs with a lower frequency.

3.6.3. Experiment 1

This experiment uses an artificial language paradigm to test whether adult participants follow the Sufficiency Principle in generalizing a rule.

3.6.3.1. Methods

3.6.3.1.1 Participants

Adult participants were undergraduates at the University of Pennsylvania. Compensation for participation in the study was in the form of course credit. A total of 25 participants were tested in the study, and each participant was tested individually with an experimenter present in the room.

3.6.3.1.2 Materials

The artificial language designed for this experiment consists of one novel subject *Vos*, five novel objects, ten novel verbs, and one novel causative morpheme *-ka*, all of which were adapted from Wonnacott, Newport, and Tanenhaus 2008. The novel items are all monosyllabic, and conform to English phonotactics. The five novel objects used are *Tom*, *Flug*, *Blerg*, *Nag*, and *Slag*. The ten novel verbs are *glim*, *shen*, *norg*, *gund*, *ern*, *loom*, *sem*, *mer*, *frag*, and *gof*.

These ten verbs were all presented in two language groups. In one language group, 5 out of the 10 verbs occurred both in a transitive causative sentence and in an intransitive sentence. In the second language group, 8 out of 10 of the verbs occurred both in a transitive causative sentence and an intransitive sentence. This means that in the 5 out of 10 group, five verbs follow the causative alternation, and in the 8 out of 10 group, eight out of ten verbs follow the rule. The remaining verbs in both language groups occurred only as intransitives.

The intransitive sentences in this language were presented to participants with only

a verb and an object in a V-O word order. The causative transitive sentences occurred in V-*ka*-S-O word order. This word order was chosen for two reasons: first, to differentiate the syntax of the language from English, and second, to front the verb to help participants learn the new vocabulary. Since the verb is crucial in this experiment, it occurred first to increase the chance that participants would learn it.

The items used in the experiment are summarized in Table 3.5. As seen in Table 3.5, each object occurs with 2 verbs. For the transitive causative forms, the same subject *Vos* was used in all instances.

Table 3.5: Novel items used in the artificial language. The subject *Vos* was heard and shown with each verb that was presented as a causative. Each of the five objects occurred with two of the ten verbs.

Verb	Causative Morpheme	Object	Subject
glim, shen	- <i>ka</i>	Flug	Vos
norg, gund	- <i>ka</i>	Blerg	Vos
flern, loom	- <i>ka</i>	Slag	Vos
sem, mer	- <i>ka</i>	Nag	Vos
frag, gof	- <i>ka</i>	Tom	Vos

Participants in the experiment were also exposed to pictures associated with the sentences. For the intransitive sentences, participants saw only the object by itself. With a causative transitive sentence, participants saw pictures of the subject *Vos* acting on the object.

A sample sentence that participants would hear in the transitive causative form is *norg-ka Blerg Vos*, where the object is depicted as being sliced into three parts (Figure 3.8). A sample intransitive sentence is *frag Tom*, where a picture depicts an object as changing color, but participants only hear the verb in the intransitive form (Figure 3.9).

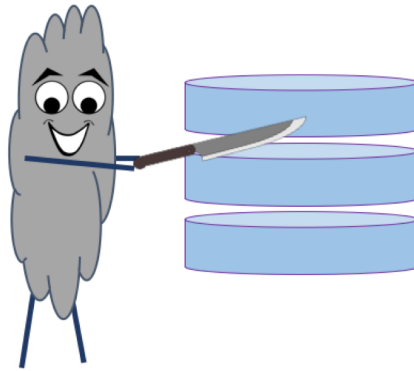


Figure 3.8: A transitive causative sentence: *norg-ka Blerg Vos* in V-ka-O-S word order. Subject is shown to act on the object and slicing it into pieces with a knife. The agent is always present in the image associated with the transitive form of the verb.

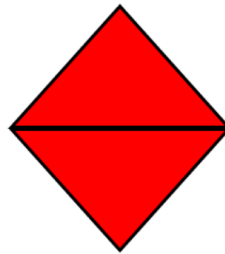


Figure 3.9: An intransitive sentence: *frag Tom* in V-O word order. There is no agent present in the image associated with the intransitive frame.

Both language groups consisted of 112 sentences in total. Out of the 112 sentences, only 42 instances of a causative form were heard in both the 5 out of 10 and the 8 out of 10 groups. The token frequency of how often the causative form was heard, or how often the intransitive form was heard, did not vary between conditions.

The frequency of the 112 items that participants were exposed to in the main task followed a Zipfian distribution for the verbs to imitate the frequency of words in natural language input. The first most frequent verb occurs 40 times, the second most frequent

verb occurs 20 times, and the five least frequent verbs occur 4 times each. The other verbs occur 12 and 8 times. This distribution is shown in the Table 3.6. The verbs used in the two language groups and their frequencies are shown in Figure 3.10.

Table 3.6: The distribution of verb frequency in the artificial language.

Verb	Frequency
Verb 1	40
Verb 2	20
Verb 3 & 4	12
Verb 5	8
Verb 6 – 10	4

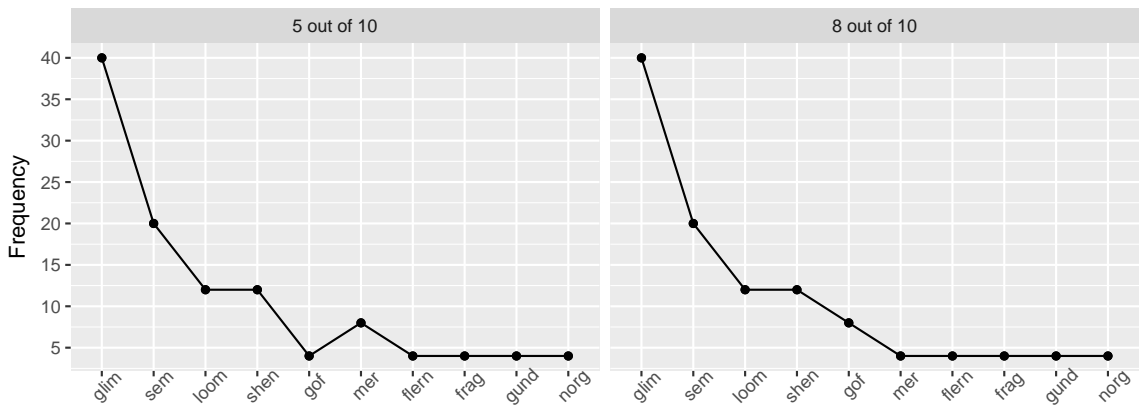


Figure 3.10: Verbs used in the 5 out of 10 and 8 out of 10 language groups with their frequencies.

3.6.3.1.3 Apparatus

The sentences were recorded by a female native speaker of American English who tried to make the sentences sound “happy” and “engaging”. Each sentence was recorded as a whole. The sound clips were all normalized to 70db in PRAAT (Boersma 2002). No text occurred with the pictures and the sound clips in this experiment.

3.6.3.1.4 Procedure

Participants were randomly assigned to either the 5 out of 10 language group, or the 8 out of 10 language group. Each participant was exposed to a training phase before the main part of the experiment. The training phase was included to familiarize participants to the experiment. Both the training and the main task of the experiment included an exposure phase and a testing phase. The exposure phase in the training phase presented participants with two novel objects, *Fleb* and *Dob*. These objects did not feature in the main experiment, and therefore, did not interfere with the task in the main experiment. Participants either saw one of two of the novel objects on the screen in each trial. When two of the same novel object occurred, participants heard the noun used with a plural marker *-po*. They were required to repeat what they heard in each trial. In the testing phase, participants were shown the object on its own to remind them of the new vocabulary. They were then shown the object as a pair, and were asked to describe the image. The elicitation task in the training phase was also designed to be used as an exclusion criteria for the experiment; however, all participants were able to complete the training task and no one was excluded.

After the training task, participants are exposed to the main task of the study. In the exposure phase of the main task, participants hear the intransitive and transitive causative forms of the novel verbs with the picture associated with it. Each verb occurs at least four times in a row to allow participants enough time to learn the pattern. In addition to the intransitive and transitive causative sentences, participants also hear and see novel objects in isolation before they are exposed to sentences containing them. This is meant to enforce the idea that the sentences consist of subjects, verbs, and objects. Participants hear no more than 6 sentences in a row before being presented with the relevant subject and object again, as a reminder. The number of times the subject and objects are heard do not vary between the two language groups. Trials are self-paced; however, the transitions from one trial to the next is facilitated by the experimenter.

After completing the exposure phase, participants were tested in an elicitation task. They were shown a prompt of the intransitive form of the sentence along with a picture of the object. Then, they were shown a picture with the subject acting on the object, and asked to describe what was happening in the picture.

3.6.3.2. Predictions

There are three main predictions of the experiment. The first prediction is regarding the generalizability of a rule. Following the Sufficiency Principle, we expect participants to generalize the rule for adding *-ka* with the agent subject only when there is enough positive evidence to do so. In this experiment, according to the Sufficiency Principle, 5 out of 10 verbs displaying a causative form does not constitute sufficient evidence for the causative rule to be extended to the other verb forms. 8 out of 10 verbs occurring with a causative form, on the other hand, is sufficient evidence for participants to assume that the rule is more general. We then expect the rate of *-ka* uses with the agent subject to be lower among predicates that were not heard with the causative form in the 5 out of 10 group with only five verbs following the rule than in the condition with eight predicates following the causativization rule. To be precise, the prediction is that participants in the 5 out of 10 group will not use the causative form with any of the predicates that did not occur in the causative frame in the exposure phase. In the 8 out of 10 group, participants are expected to use the causative form even with verbs that were not presented in the causative frame in the exposure phase. This prediction is illustrated in Figure 3.11.

The second prediction is in regards to theories of acquisition that invoke entrenchment as a learning mechanism. Here, as the verbs in the exposure phase follow a Zipfian distribution, the entrenchment theory of learning predicts a distinction in the use of purely intransitive verbs based on their frequency (e.g., Ambridge et al. 2008). In the 8 out of 10 condition, entrenchment predicts that participants are less willing to use the causative form with a purely intransitive verb that occurs with a relatively high frequency than a

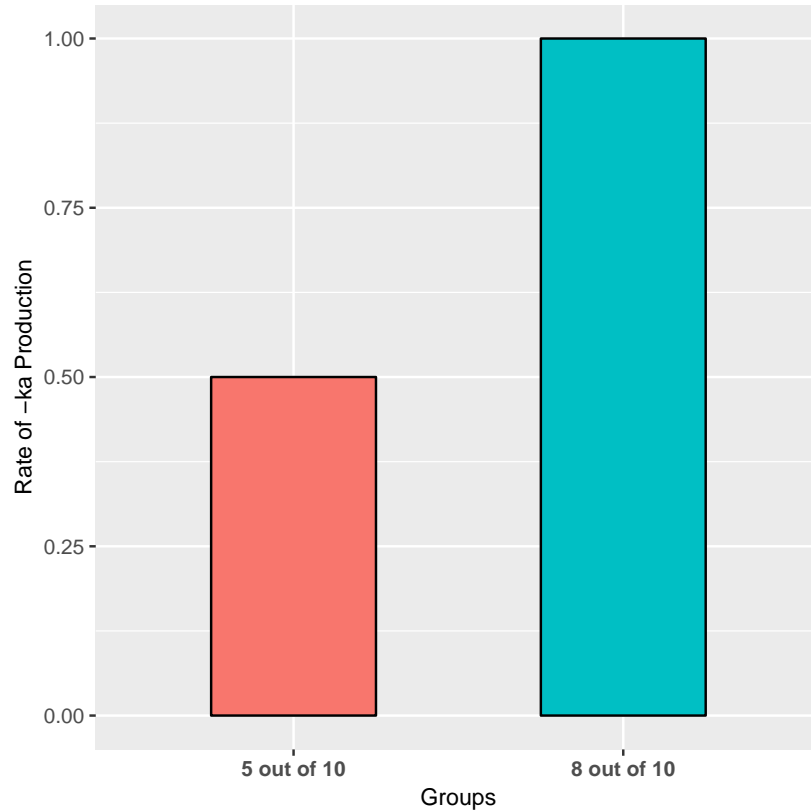


Figure 3.11: Prediction for the 5 out of 10 language group with 5 *-ka* verbs and 5 no *-ka* verbs and 8 out of 10 language group with 8 *-ka* verbs and 2 verbs without *-ka*.

verb that occurs with a lower frequency. If participants only generalize in the 8 out of 10 group, then we may not see these effects for the 5 out of 10 language group. However, entrenchment theories do not predict a difference between the two groups; rather, they predict a difference between the verbs within each group based on their frequencies. In the 8 out of 10 group, the verb *sem* occurs 20 times as a pure intransitive in the exposure phase, while the other pure intransitive verb *gof* occurs only 8 times. If participants do not treat *sem* differently from *gof*, then we find no support for entrenchment. In the 5 out of 10 group, *loom* occurs 12 times and is the pure intransitive verb with the highest frequency. The other pure intransitive verbs occur 4 times each.

One point that should be recalled regarding entrenchment-based approaches is that the data are often obtained through participants' ratings on a likert scale for the ungrammatical forms of various verbs. These ratings are directly translated into verb grammat-

icality judgments, when there may in fact be an effect in the ratings that are unrelated to grammaticality. Thus, the ratings found in these approaches do not inform our understanding of the generalizations the learner is making. Moreover, any frequency effects that may be obtained through the rating tasks in entrenchment-based approaches are impertinent to the learner's decision to generalize a rule. In these experiments, I aim to test the learner's ability to generalize. However, the design of the experiments also allow for the testing of frequency effects in generalizing from the input.

Another prediction made here is based on previous studies (e.g., Schuler, Yang, and Newport 2016) is that adults probability match instead of following the Sufficiency Principle or showing entrenchment effects, unlike children. This pattern may be due to the fact that adults employ different learning mechanisms than children. If adults in this experiment are indeed probability matching, we expect them to produce the causative form about a third of the time, which is their frequency in the input. If participants probability match, the number of predicted *-ka* tokens would be the same in both language groups, as the number of causative tokens in both groups was the same in the exposure phase. This prediction is illustrated in Figure 3.12.

Finally, it should also be noted that if adults in this experiment do not follow the Sufficiency Principle, then that does not indicate that the Sufficiency Principle does not hold for child language acquisition; it only indicates that adults do not follow the Sufficiency Principle. In this case, the results could indicate that the Sufficiency Principle only applies to child language learning. This experiment, therefore, must crucially also be tested on child participants, which is done in Experiment 3.

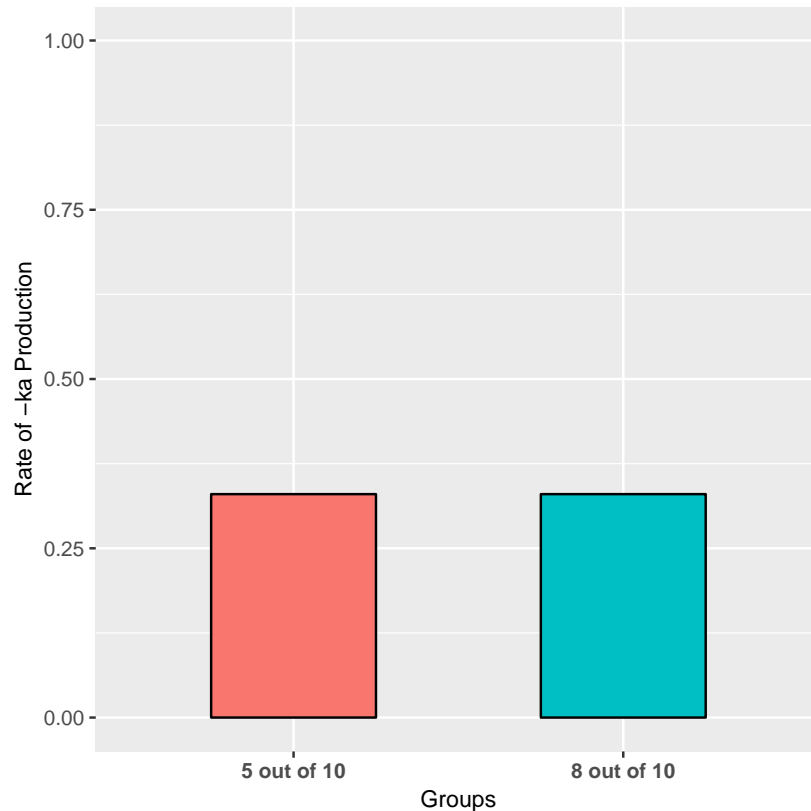


Figure 3.12: Participants are predicted to produce *-ka* about a third of the time in both conditions if they probability match the token frequency of *-ka* in the exposure phase. The percentage of *-ka* production reflected in the graph is the total number of causatives produced out of the token occurrences of the 10 novel verbs.

3.6.3.3. Results

The results are summarized in Figure 3.13. The results show that participants overgeneralize in both the 8 out of 10 group as well as the 5 out of 10 group. Only one participant in the 5 out of 10 group did not overgeneralize the *-ka* morphology; however, this participant did produce the subject *Vos*. As the causative morphology is a key component of the causative alternation, the responses for this participant were not scored as producing causatives. A Fisher's Exact test was conducted to test whether the responses between the two groups was significantly different. Results from Fisher's Exact test indicate that there was no significance between the responses in the two groups ($p = 0.3$). If the participant who produced the subject without the causative morpheme was included in the calcu-

lations, there would be even less of a difference in terms of the number of *ka* responses between the two groups.

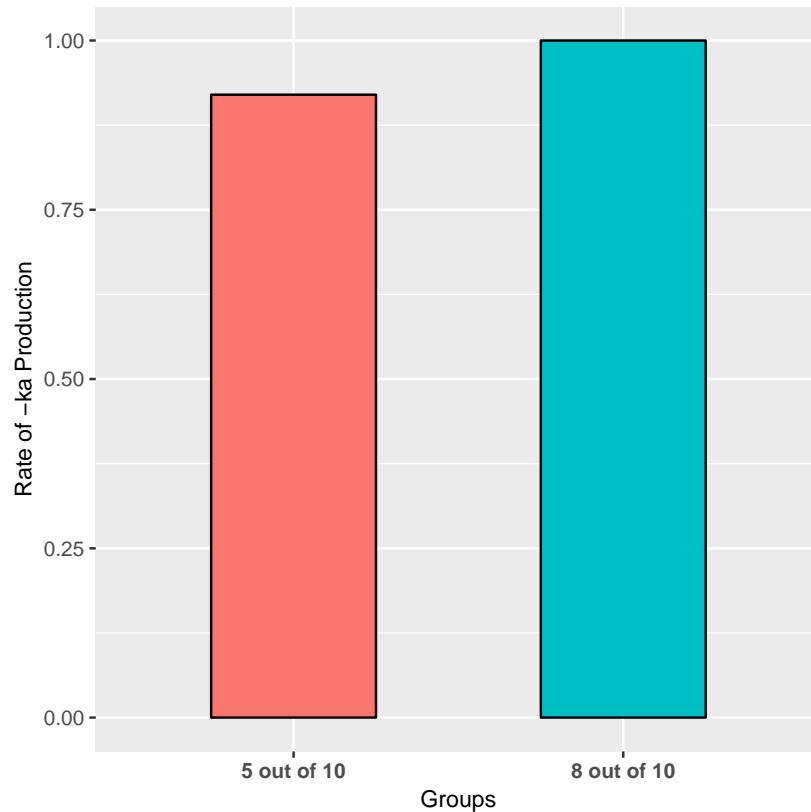


Figure 3.13: Rate of *-ka* production across the two language groups.

Since the participants produced the causative form categorically for all verbs within a language group, the results also indicate that there is no difference in the number of *-ka* responses with more frequent verbs versus less frequent verbs. In other words, the results do not show an entrenchment effect in the *-ka* responses of either language group. Instead, participants consistently produced the causative form of the verb in the testing phase. All but one of the participants produced the *-ka* morpheme when tested, and even this participant produced the subject.

3.6.3.4. Discussion of the Results

The results of Experiment 1 suggest that the causativization rule is still learnable by adults even when the number of items that follow the rule is below the threshold determined by the Sufficiency Principle. However, there are several reasons why the results found here do not strictly provide evidence to reject the Sufficiency Principle.

One reason for why there was no difference in responses between the two language groups could stem from the fact that adults and children use different learning mechanisms. The Sufficiency Principle is a model of child language acquisition, and therefore, may not apply to adults. This possibility can be tested by running the same paradigm on children, which is done Experiment 3.

Another reason for why the results may not necessarily disconfirm the Sufficiency Principle comes from the fact that there was a strong correlation between the presence of a subject in the images, and the use of the causative form. In Experiment 1, whenever participants saw the subject *Vos*, they heard the novel verb in the transitive causative frame. Participants never saw the subject *Vos* when they heard the novel verb in the intransitive frame. Therefore, when they were tested on the verbs that they only heard as intransitives, they might have been tempted to use the transitive causative form due to the presence of the subject. Therefore, in Experiment 2, the subject was eliminated from the language. Although participants still see an agent acting on the object in the images, the elimination of the subject from the actual sentences is expected to reduce the strong correlation between the causative marker and the subject.

In Experiment 1, the responses found for adult participants may have also resulted due to a last-resort strategy. It is possible that the experiment required a high memory load, and thus, participants were unable to learn the individual properties of the items. As a result, they may have overgeneralized if they found it too difficult to recall the novel sentences heard in the experiment. However, it should be noted that previous experiments often find that adult participants probability match in these cases (e.g., Schuler et

al. 2018). Given previous results, it is still surprising that participants in Experiment 1 chose to overgeneralize as opposed to using a different last-resort strategy such as probability matching.

3.6.4. Experiment 2

Initially, the same paradigm used in Experiment 1 was to be used to test children in Experiment 2. However, the paradigm used in Experiment 1 proved to be too complex for children. Thus, some changes to Experiment 1 were made in order to make it more suitable for children. The aim of Experiment 2 is the same as in Experiment 1, which is to test the threshold of generalization for adults.

Experiment 2 tests adult participants. Although adults in Experiment 1 overgeneralized in both language groups, a potential explanation for this result could be due to a last-resort strategy stemming from the complexity of the experiment. Experiment 2 was designed with a simpler task, which allows us to address the question of generalization with less confounding factors in place.

3.6.4.1. Methods

3.6.4.1.1 Participants

Adult participants were undergraduates at the University of Pennsylvania. Compensation for participation in the study was in the form of course credit. There were a total of 20 participants; 10 subjects were run in each condition. Each subject was tested individually with an experimenter present in the room.

3.6.4.1.2 Materials

Several measures were taken to simplify Experiment 2. The artificial language created for Experiment 1 was reused in Experiment 2; however, one change made to Experiment 2 was that the subject *Vos* was removed from the language. This resulted in a Verb-Object word

order for intransitive sentences and a Verb-*ka*-Object word order for transitive sentences. Thus, the artificial language still resembled the causative alternation in natural language, while also allowing for minimal difference between the two sentence types.

Second, in addition to pictures and recordings, participants were also shown the language in written form as additional visual support in Experiment 2. The text was added in order to facilitate learning of the language by breaking it into its component parts. The pilot data also confirmed that participants found the task easier when provided with text that indicated the separate words used in the auditory stimuli. The rest of the stimuli used were the same as in Experiment 1.

3.6.4.1.3 Procedure

The language groups were the same as in Experiment 1, and the tasks still consisted of an exposure phase and a testing phase using an elicitation task. However, the duration of the task was considerably shortened for Experiment 2. During the main task of the experiment, participants only heard the 112 sentences; they were not exposed to any individual items, as opposed to in Experiment 1.

Additionally, the training phase in Experiment 1 was eliminated, and replaced by a noun matching game instead. There were two rounds in the noun matching game. In the first round, participants see an object in the middle of the screen, and are asked to match this object with one of two choices at the bottom of the screen. They are asked to repeat the name of the object shown in each trial. In the second round of the matching game, there is no object prompt in the middle of the screen. Participants hear the name of an object, and have to select one of two options on the screen; i.e., the correct object that matches with the novel word heard in the trial. The matching game was introduced in order to help participants learn the names of the objects used in the language. The game was meant to help participants recognize that the sentences used in the exposure phase consisted of nouns and verbs. The matching game was deemed more helpful than the

previously used training task in order to get participants to learn the new language.

Another change that was introduced in this study is the addition of sticker breaks. There were 6 sticker breaks randomly interspersed throughout the experiment. The sticker breaks were included to motivate participants to complete the study, and to give them a chance to take a break from the tasks in the experiment.

3.6.4.2. Predictions

The predictions of this experiment are the same as in Experiment 1 in Section 3.6.3.2 (cf. Figures 3.11 and 3.12).

3.6.4.3. Results

The results are presented in Figure 3.14. Just as in Experiment 1, participants in Experiment 2 overgeneralized in both the 8 out of 10 group as well as the 5 out of 10 group. Each of the 10 participants in both groups produced *-ka* for each novel verb item tested. There was no difference between the responses of the participants in the two language groups.

The results of Experiment 2 also indicate that participants are not more likely to assume that the causative marker is ungrammatical for a more frequent verb. As in Experiment 1, Experiment 2 also found no effect of verb frequency.

3.6.4.4. Discussion of the Results

The results suggest that adults, at least in this paradigm, do not follow the Sufficiency Principle, nor are they simply probability matching. Instead, they have learned the causative alternation rule in both the 5 out of 10 and 8 out of 10 groups. The adult responses indicate that even 5 out of 10 verbs following the causative alternation rule were sufficient for generalization.

There are two potential interpretations of the results from Experiments 1 and 2. The first interpretation is that 5 out of 10 items showing a given property is enough for the

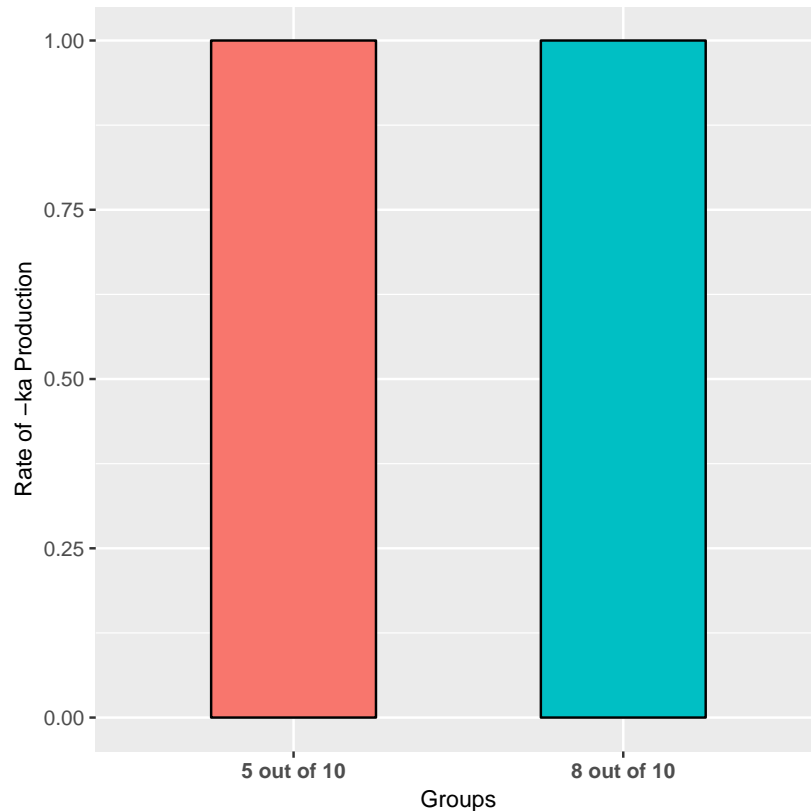


Figure 3.14: The results for Experiment 2 show that participants overgeneralized the causative alternation rule in both the 5 out of 10 and the 8 out of 10 condition. Each participant used the causative form for each response tested.

participants to overgeneralize. If adults have indeed successfully learned the causative alternation rule with 50% of the items showing the pattern, then the results are an indication that adults behave distinctly from the child learner, as predicted by the Sufficiency Principle.

Moreover, the second prediction of adults probability matching was based on previous studies that found that adult participants probability match in similar paradigms (e.g., Schuler, Yang, and Newport 2016). This pattern was not found in this study. The fact that different results were found in this study as opposed to previous studies testing the Tolerance Principle leads us to probe deeper into the aspects of the studies that differed. For instance, Schuler et al. 2016 test positive exceptions in the experiments reported therein; i.e., participants have positive evidence for items that do not follow the rule. In the exper-

iments presented here, negative exceptions are tested; participants do not hear whether an item is an exception to the rule. Therefore, the results of Experiments 1 and 2 suggest that adults may treat negative and positive exceptions to the rule differently.

Finally, the results also show that no effects of entrenchment were found. The most frequent pure intransitive verb occurred 20 times in the 8 out of 10 condition. The low frequency intransitive verb occurred only 4 times in the exposure phase. The results show that participants were not more likely to causativize the low frequency intransitive verb as opposed to the high frequency verb.

3.6.5. Experiment 3

Experiments 1 and 2 tested adult participants to see when they would generalize a rule. In Experiment 3, child participants were tested instead. The Sufficiency Principle is a model of child language acquisition, and therefore, in order to test whether it can be used to model child language learning, it must be tested on children.

3.6.5.1. Methods

3.6.5.1.1 Participants

Participants were 26 children between the ages of 5;9 years and 9;3 years. The mean age of the children was around 7;2 years with a standard deviation of 11.4 months. 13 children were run in each condition with an experimenter present in the room. All of the children tested had corrected to normal vision. Children were given a book for their participation in the study, and their caretakers were given \$5 to compensate for their travel to the testing location.

3.6.5.1.2 Materials

The stimuli used in this experiment were the same as in Experiment 2.

3.6.5.1.3 Procedure

The procedure of Experiment 3 did not vary from Experiment 2.

3.6.5.2. Predictions

This experiment has two main predictions, which are similar to those presented for Experiments 1 and 2. The main difference between Experiments 1 and 2, and Experiment 3 here is that children, unlike adults, are expected to follow the Sufficiency Principle and not probability match. As this experiment tests children instead of adults, the results will directly address the predictions made by the Sufficiency Principle.

The first prediction is regarding the generalizability of the causative alternation rule. The two language groups differ in the number of novel verbs that follow the rule. If participants are adhering to the Sufficiency Principle, we expect overgeneralization to occur in the language group where 8 out of 10 novel verbs show the causative alternation. The Sufficiency Principle predicts that participants should not overgeneralize when they only hear 5 out of 10 novel verbs that show the alternation. Hence, we expect the rate of *-ka* production to be greater in the 8 out of 10 group than the 5 out of 10 group, as shown in Figure 3.11.

As the novel verbs in the exposure phase occur in a Zipfian distribution, the experiment also makes a prediction regarding an entrenchment-based approach of language learning. Under an entrenchment approach, high frequency verbs that do not occur in a transitive causative form are predicted to be more likely to be used as intransitives than low frequency verbs. A high frequency pure intransitive verb is said to become entrenched in the linguistic forms in which it is used over time. If verbs do become entrenched in the forms they are used, we would expect participants to show a difference between high frequency and low frequency intransitives. Since the Sufficiency Principle predicts no overgeneralization in the 5 out of 10 group, the 8 out of 10 group is where we expect to find an entrenchment effect, if any effect is found. It should also be noted here

that an entrenchment-based approach does not predict a difference between the two language groups. The only difference predicted is between verbs with different frequencies within a language group.

The two verbs that do not occur in a causative frame in the 8 out of 10 group vary in terms of their frequency. The high frequency novel verb occurs 20 times as a pure intransitive, whereas the low frequency novel verb occurs 8 times. Participants are predicted to be more likely to recall that they heard the high frequency novel verb as an intransitive than the low frequency novel verb. If there is a difference in the rate of *-ka* production between the more frequent and the less frequent intransitive verbs in the 8 out of 10 condition, then we find evidence for an entrenchment effect.

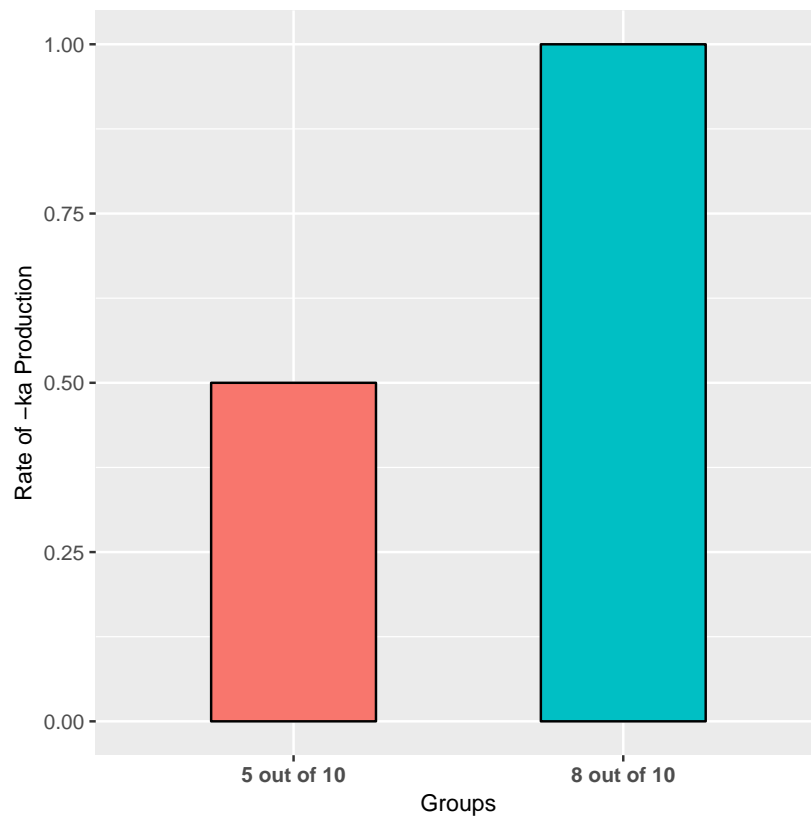


Figure 3.15: Predicted results for Experiment 3. Prediction for the 5 out of 10 language group with 5 *-ka* verbs and 5 no *-ka* verbs and 8 out of 10 language group with 8 *-ka* verbs and 2 verbs without *-ka*. Participants are only expected to produce *-ka* for the verbs that were presented as a transitive causative. For the 8 out of 10 group, participants are expected to produce *-ka* across the board.

Finally, we do not expect that child participants will probability match their *-ka* production with the token of frequency of *-ka* heard in the exposure phase. Adults may not have the same learning mechanism as children, and may, therefore, show this behavior of probability matching. Children on the other hand are predicted to follow the Sufficiency Principle.

3.6.5.3. Results

Participant responses were excluded from the analysis if the wrong verb was produced, but were included if they mispronounced the object or said the name of a different object. This resulted in the exclusion of 11 out of 260 responses from the analysis. Additionally, one participant had to be excluded from the 5 out of 10 group because they produced the wrong verb in more than 50% of the trials.

The results are presented in Figure 3.16. A Fisher's Exact test was conducted on the child responses in the two language groups, and the test shows a significant difference in the responses between the two groups ($p < 0.001$). Participants in the 5 out of 10 group produced significantly less causatives than participants in the 8 out of 10 language group. This is illustrated in Figure 3.16. These results were obtained by taking each participant's response rate into consideration, as opposed to collapsing all of the participants responses and treating the ratio as one.

The responses of each participant grouped by language can be seen in Figure 3.17. In the 5 out of 10 group, participants generally either overgeneralized completely or did not produce any causative forms at all, even for the verbs that they heard in the causative frame in the exposure phase. In the 8 out of 10 group, all but one participant overgeneralized the causative form to at least one of the two verbs that occurred only as an intransitive in the exposure phase. This is in contrast to 7 participants who did not robustly overgeneralize in the 5 out of 10 group.

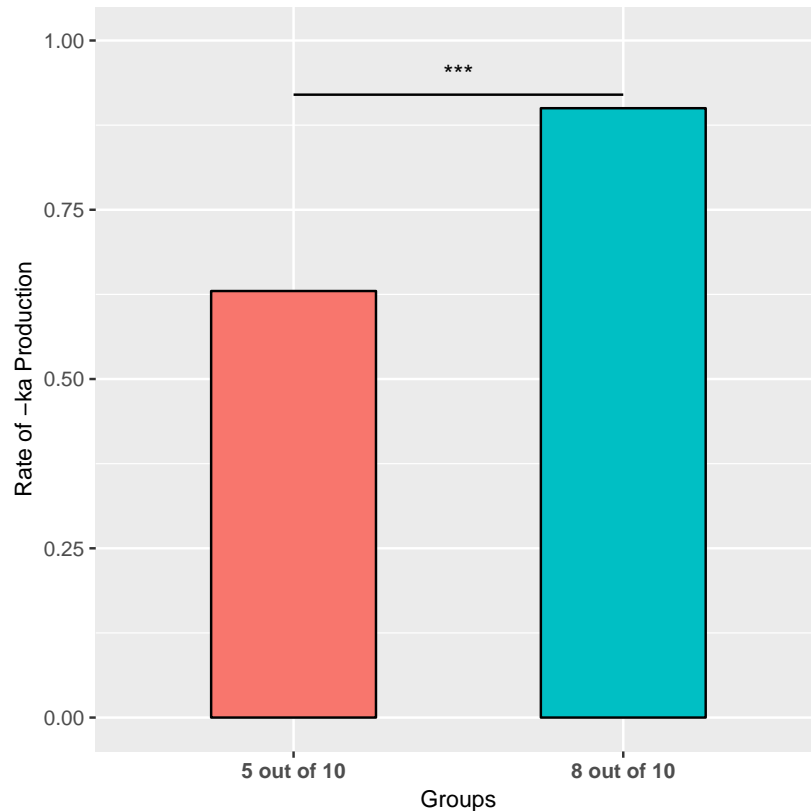


Figure 3.16: Summary of children’s responses in Experiment 3. Children in the 5 out of 10 language group produced significantly less causatives than participants in the 8 out of 10 language group.

The results also show that participant responses were mostly categorical, indicating that verb frequency did not influence generalization. In the 8 out of 10 group, the more frequent intransitive verb occurred 20 times in total, and the less frequent verb occurred 8 times overall. Fisher’s Exact test reveals no difference between these verbs ($p = 0.593$). The number of transitive causative productions with *-ka* by verb are summarized in Figure 3.18. Furthermore, the graphs also illustrate that no significant effect of frequency was found in the causative usage of verbs that were presented as intransitives in 5 out of 10 group. Even for the intransitive verbs that had the most numerical difference in causative usage, the effect did not reach significance (*loom* (12 times) versus *gof* (4 times) $p = 1$). The proportion of use of a causative form for verbs that were only presented to participants as intransitives are illustrated in Figure 3.19.

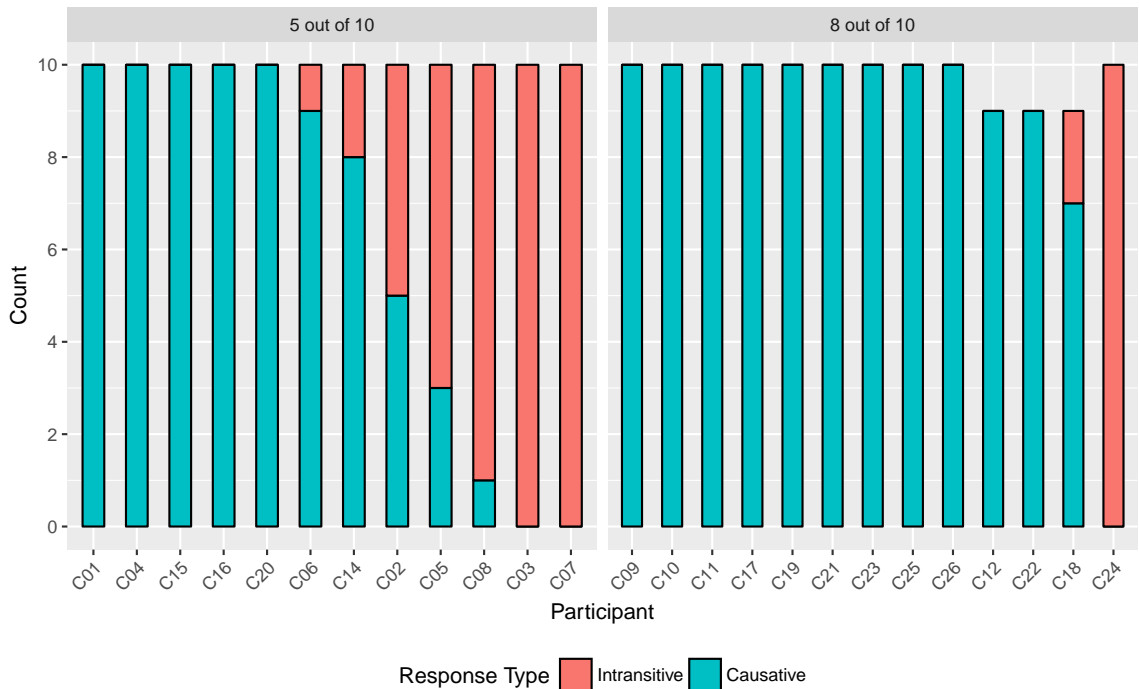


Figure 3.17: Results from testing children in Experiment 3. Number of causative and intransitive forms produced by participants grouped by the language they were tested in. All but one participant overgeneralized to some degree in the 8 out of 10 group, whereas 6 participants did not overgeneralize in the 5 out of 10 group. Almost all participants either overgeneralized completely or failed to produce to causative marker at all.

3.6.5.4. Discussion of the Results

The results from Experiment 3 show that in the 8 out of 10 group, all but one of the 13 children overgeneralized the rule. This is in stark contrast to the responses of the participants in the 5 out of 10 group. This indicates that 8 out of 10 verbs undergoing the causative alternation is sufficient evidence for participants to overgeneralize it even to verbs that only occurred as an intransitive, but 5 out of 10 verbs constitutes insufficient evidence.

The significant difference between the responses of the two groups of children indicates that the amount of positive evidence present in the two language groups has an effect on rule learning for children. The child participant responses are in line with our prediction for the data if the children were following the Sufficiency Principle. Therefore,

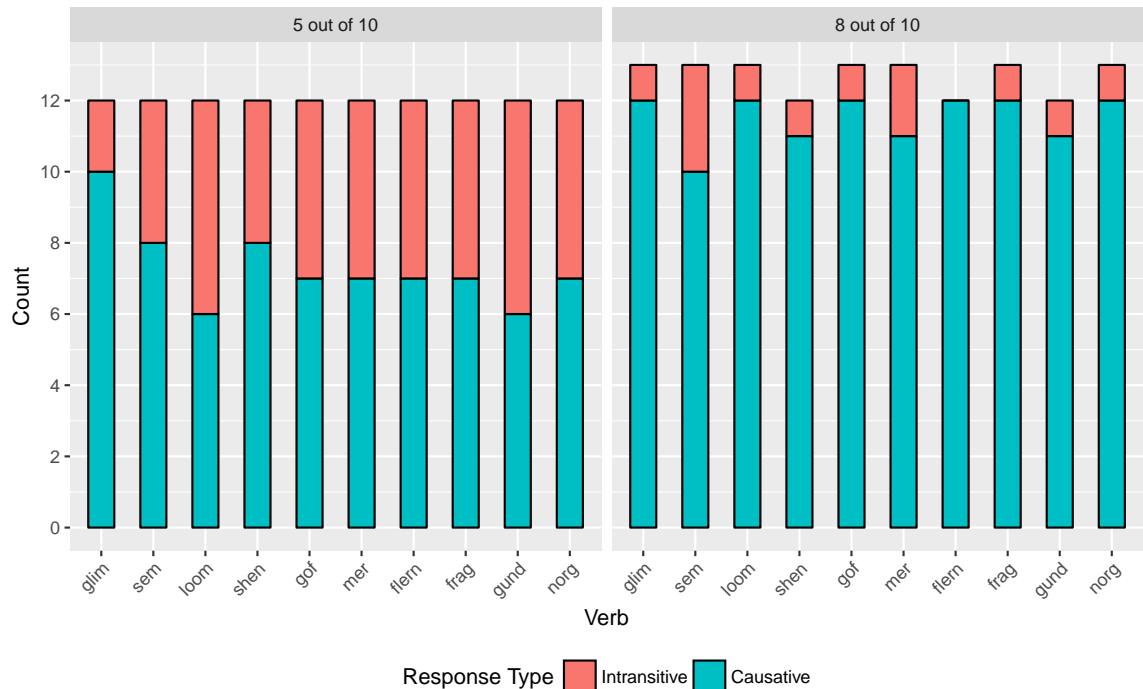


Figure 3.18: Results from testing children in Experiment 3. Number of causative forms produced by participants with each verb in the 5 out of 10 and 8 out of 10 language groups. The items on the x-axis are ordered according to their frequencies, with higher frequency items on the left. No significant effects of frequency were found in the production of causatives.

the data are compatible with a learning model that uses the threshold determined by the Sufficiency Principle to generalize a rule.

Additionally, the results do not show an effect of verb argument structure entrenchment. Participants generally either overgeneralized the causative form to every verb, or they failed to use the causative form with any of the verbs. As discussed above, this may be due to the fact that the children did not remember which verbs take a causative form, and therefore they did not produce any causatives. Any overgeneralization in the 5 out of 10 group could also be due to the fact that children only learned the more frequent verbs, which did take a causative form. Only 6 participants did not perform at ceiling or floor level. Out of these 6 participants, one in the 5 out of 10 group, participant C02, actually followed the Sufficiency Principle perfectly. This participant only produced the

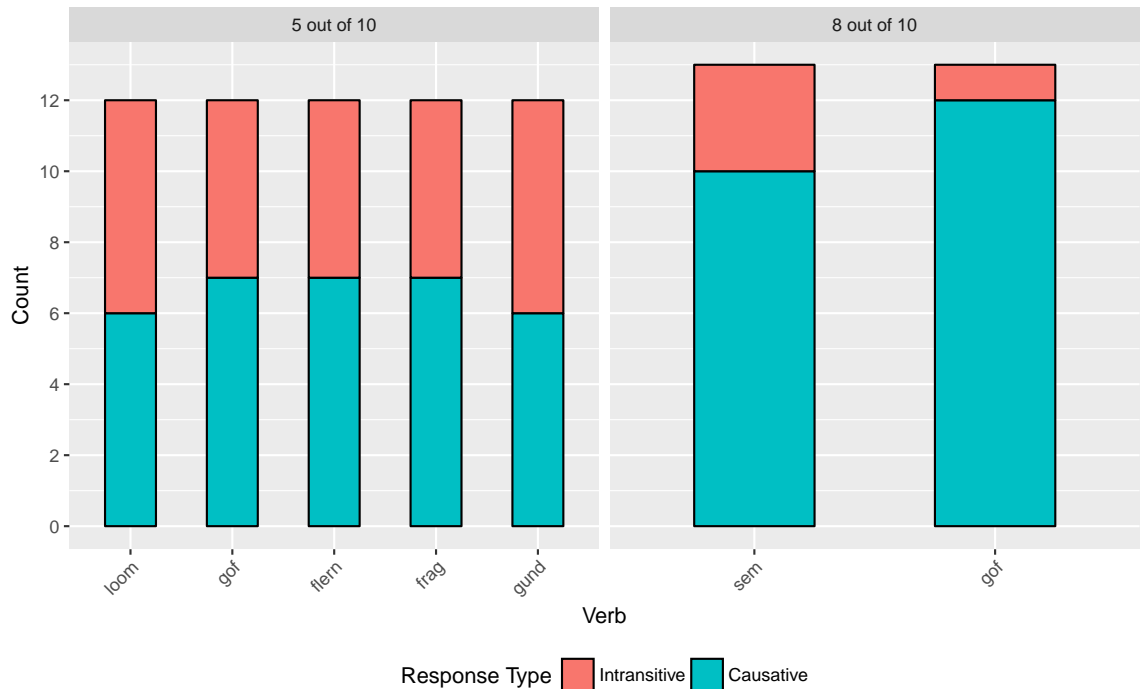


Figure 3.19: Results from testing children in Experiment 3. Number of causative forms produced by participants with each verb presented as an intransitive in the 5 out of 10 and 8 out of 10 language groups. The items on the x-axis are ordered according to their frequencies, with higher frequency items on the left. No significant effects of frequency were found between the intransitive verbs in either group.

causative form for verbs that occurred as a causative in the input. This leaves us with 5 out of the 25 participants examined who did not behave categorically. Out of these 5, 4 were in the 5 out of 10 language group. Their results are compatible with the prediction of the Sufficiency Principle. Thus, entrenchment does not account for the pattern we find in the responses of these participants; however, it should be noted that the number of instances in this experiment may not be large enough to find an entrenchment effect. Regardless, a difference between the two language groups was found for child participants which cannot be accounted for under a solely entrenchment-based approach.

It should be noted here that although the data are compatible with the hypothesis that children are following the Sufficiency Principle when generalizing the rules of their language, there were large individual differences in the responses. In the 5 out of 10 group,

we see that 6 out of 13 participants did not overgeneralize the causative alternation rule. In fact, 5 out of these 6 participants omitted the causative marker even when they had heard the novel verb in the causative transitive frame in the exposure phase. Given our prediction, this result requires further discussion. One possible reason for omission of the *-ka* marker is that participants may have been aware that some of the verbs could undergo the causative alternation, but they couldn't remember which ones. As a result, children resorted to using the intransitive form of the verb across the board. For these participants, because overgeneralization was not an option, they used the intransitive form when they could not remember which novel verbs took the causative form. This result is not unexpected given that the experiment took place in one day, and that participants had roughly fifteen minutes of exposure to the language on average. Therefore, we may expect different results if the experiment took place over multiple days, and if participants had more time to remember the argument structure of each individual verb heard in the exposure phase.

Furthermore, 7 out of the 12 participants did overgeneralize in the 5 out of 10 condition to some degree. Under our account, no participant is expected to generalize in this language group due to insufficient positive evidence. However, this result may have been obtained due to participants only recalling the more frequent verbs in the language. In the 5 out of 10 language group, many of the more frequent verbs did in fact undergo the alternation, while the less frequent verbs occurred as pure intransitives. In this group, 4 out of 5 least frequent verbs occurred only as an intransitive, which means that even if participants only learned 9 verbs instead of 10, they would have stored 5 out of 9 verbs in the language as undergoing the alternation. According to the Sufficiency Principle, for a class with 9 members, 5 is sufficient evidence to generalize a rule ($9 - 9/\ln(9) = 5$). This kind of result is indeed what was found by Schuler et al. 2016 in testing the Tolerance Principle. Schuler et al. 2016 also included a rating task in their experiment, which confirmed that overgeneralization in the condition with insufficient evidence sometimes took

place when participants only learned a subset of the verbs. However, due to the length of the present study, a rating task was not included in Experiment 3.

3.6.6. General Discussion

Experiments 1, 2, and 3 set out to test whether adults and children adhere to the Sufficiency Principle in learning the rules of their language. In Experiment 3, 12 children were exposed to an artificial language where 5 out of 10 novel verbs underwent the causative alternation. 13 other children were exposed to the same artificial language, but instead, 8 out of 10 of the novel verbs underwent the causative alternation. The prediction here was that children's responses would vary depending on the version of the language they were exposed to. Indeed, we see that children did not overgeneralize in the 5 out of 10 group when the number of verbs that participated in the rule were below the threshold of the Sufficiency Principle ($10 - 10/\ln 10 = 5.7$). Children in the 8 out of 10 group overgeneralized robustly.

Another aspect that is relevant for understanding how children generalize rules is that of retreat from overgeneralization. In the 5 out of 10 language group in Experiment 3, children at some point would have likely generalized the causative rule because the verbs that only occur as intransitives were infrequent in the exposure phase. 4 out of the 5 verbs that did not alternate, were only heard 4 times each. Thus, children likely formed a rule that would prompt them to overgeneralize at some point in the experiment, and eventually retreated from that generalization.

The results of experiments 1 and 2 conducted on adults are especially surprising given that previous studies found that adults tend to probability match in similar experiments (e.g., Schuler et al. 2016). Moreover, Wannacott et al. 2008 also conducted a similar study where they found that adults did not overgeneralize when there was insufficient evidence in the input. There are some potential explanations for why there is a difference between the experiments conducted here and previous studies. One possible reason is that unlike

children, adults are willing to generalize when at least half of the items follow a rule. In Wannacott et al. 2008's study, they tested 4 out of 12 verbs that did not follow a rule, which is less than 50%. In this case, adults were able to recall the verb argument structure correctly. In contrast, in experiments 1 and 2, adults in both conditions were shown the causative form for at least 50% of the verb types in both language groups. A second reason for the difference in the adult responses in Wannacott et al. 2008's study and the present study could stem from the methodology employed. For instance, participants in the Wannacott et al. 2008 study were tested over multiple days instead of within a single day.

Another reason for the difference between the adult and child responses is that adults may employ a potential explicit learning mechanism; i.e., they may rely on explicit correlations or instruction in their environment as opposed to the implicit structural cues in the language (e.g., Reber 1976). In this case, adults could have formed a strong correlation between the causative form and the presence of an agent acting on the object. Children, on the other hand, may be learning the implicit structural generalizations. The adult results we see here require further investigation to understand how they differ from the typically observed adult pattern of probability matching, and why adults overgeneralize consistently in the experiments reported here. Thus, the results found in Experiments 1 and 2 have implications for second language learning

A third reason for why we obtained these results could also stem from a task effect. In the exposure phase, participants were first shown the intransitive form of the novel verb with the image associated with that particular intransitive form. When tested, participants were presented with a different image, which may have biased them to produce a response that is distinct from the sentence they had just heard. The only other option for the adult participant, in that case, is the causative form. Children might not be pragmatically biased in this manner, and therefore, the results obtained for child participants would not show robust overgeneralization in both conditions.

Throughout this chapter, I have also argued against indirect negative evidence approaches to account for the acquisition of the causative alternation. In particular, it was argued that entrenchment approaches cannot account for the errors that children produce when learning the causative rule. The experiments presented in this section find no entrenchment effects for both adult and child populations. Thus, the results of the experiments further support the argument that entrenchment effects cannot be used to fully account for the acquisition of causatives. Instead, the results support learning from positive evidence in the input.

Chapter 4

The Acquisition of Passives

4.1. Introduction

In Chapters 2 and 3, I discussed the acquisition of raising verbs and unaccusative verbs, which share the core property of not introducing an external argument. In this chapter, I discuss the case of passives, which are also characterized by the lack of a syntactically projected external argument. Here, I show that the acquisition of passives is easily accounted for in the framework presented in this dissertation.

I demonstrate how the acquisition of passives can be analyzed under the Active Mapping Model. I argue that a crucial aspect of the model, which is the independence of structural and conceptual cues, derives the developmental facts observed for the acquisition of the passive construction. In this chapter, I investigate the acquisition of passives in English and Sesotho, and show that the construction is productive for the child learning these languages.

This chapter addresses previous claims made in the literature regarding passives in child grammar. There is much debate regarding when children acquire passives (Allen and Crago 1996; Borer and Wexler 1987; Demuth 1989; Driva and Terzi 2007; Lempert 1990; Pierce 1992; Pye and Poz 1988). For instance, passives in Inuktitut are argued to be acquired early (e.g., Allen and Crago 1996), but it has been claimed for children learning Greek that they acquire passives late (e.g., Driva and Terzi 2007). Similarly, English-learning children are said to acquire passives late (e.g., Borer and Wexler 1987), whereas

children learning Sesotho are shown to have full command of the construction by age 3 (Demuth 1989, Demuth et al. 2010). In this chapter, I show that English passives are productive in the input, and that the passive construction is productive for the learner early on, given the child's vocabulary and verb classes.

This chapter is structured as follows. First, I present an overview of previous theories and findings proposed for the acquisition of passives. Next, in Section 4.3, I present the analysis of the acquisition of passives argued for in this chapter. In Section 4.4, I show that the proposed analysis holds for English by conducting two corpus studies. I then discuss the acquisition of Sesotho passives in Section 4.5, and compare the Sesotho passives in the linguistic input to English passives in the input.

4.2. Background

Passives have been a consistent subject of research in the field of language acquisition (Ambridge, Bidgood, Pine, Rowland, and Freudenthal 2016; Bever 1970; Borer and Wexler 1987; Crawford 2012; De Villiers 1985; Demuth 1989; Pinker, Lebeaux, and Frost 1987; i.a.) for two prominent reasons: the acquisition of passives involve *A*-chains, which are said to be acquired late under the maturation hypothesis (Borer and Wexler 1987; Hirsch and Wexler 2006; Terzi and Wexler 2002; i.a.), and second, it has been claimed that the semantics of a verb affect the learnability of passives (e.g., Pinker 1989). Actional verbs have been shown to be acquired earlier than non-actional verbs. These non-actional verbs often overlap greatly with the class of psychological verbs (Maratsos, Fox, Becker, and Chalkley 1985; Pinker, Lebeaux, and Frost 1987; i.a.). Examples of non-actional and actional passives are shown in (45). This asymmetry between the acquisition of actional and non-actional passives is curious because even though young children are shown to comprehend actional passives, Maratsos et al. 1985 show that non-actional passives are not acquired until the child's school years. In this chapter, I provide an analysis of the

acquisition of passives that also accounts for this asymmetry.

- (45) a. The boy was pushed by all his friends. [Actional passive]
b. The boy was liked by all his friends. [Non-actional passive]

A majority of the work done on the acquisition of passives claims that the construction is acquired relatively late in English (e.g., Borer and Wexler 1987). A maturational hypothesis of language acquisition is often invoked as a way of accounting for that distinction, as this theory argues that certain grammatical processes, such as A-chains, do not biologically develop until a later age. This line of work argues that children under the age of 6 rarely produce passives, and that any passives found under that age are short passives; i.e., passives without a *by*-phrase. These passives are then claimed to be adjectival passives, and not verbal passives, where there is no formation of an A-chain.^{1,2} Examples of the different kinds of passives considered in this chapter are illustrated in (46). The maturational hypothesis also attempts to explain the asymmetry between the learnability of passives of actional versus non-actional verbs by arguing that non-actional verbs do not make good adjectival passives, and are therefore, not learned earlier. I argue against these claims in this chapter.

- (46) a. The door was open. [Adjectival passive]
b. The door was opened. [Verbal passive]
c. The building was demolished. [Short passive]
d. The building was demolished by bulldozers. [Long passive]

It should be noted here that the argument that children learning English have trouble with passives also stems from the low frequency of passives with *by*-phrases in the input

¹It should also be noted here that once we assume the VP-internal subject hypothesis (McNally 1992), all sentences that contain a subject necessarily have an A-chain. However, I leave this point aside for the purpose of the argument.

²I also leave aside here the question of children's production of get-passives, which have been observed as early as age 3. The reader is referred to Fox and Grodzinsky 1998 for a discussion on how the production of get-passives is evidence against the maturation hypothesis.

(Borer and Wexler 1987; Hirsch and Wexler 2006). These *by*-phrases are argued to be able to distinguish between adjectival and verbal passives with certainty. The lack of *by*-phrases with passives in child data has been claimed as evidence against the idea that children have mastered the passive construction, as the short passives are claimed to be adjectival passives that do not involve A-chains.

There is, however, reason to suspect that the claim that children do not produce full passives early on is inaccurate. Examining child production data in CHILDES reveals several long passive constructions with *by*-phrases produced before the age of 4, contrary to previous claims (e.g., Hirsch and Wexler 2006). Some passives produced by children are illustrated in (47).

- (47)
- a. Sammy was hit by a car (2;7, Laura, Braunwald corpus)
 - b. No a monarch butterfly was killed by a bird (3;7, Abe, Kuczaj corpus)
 - c. Oh Dad I heard that man say he was never bitten by that (3;8, Abe, Kuczaj corpus)
 - d. And the dinosaur was eaten by the alligator (3;9, Abe, Kuczaj corpus)
 - e. Your father was killed by a &trecher of mine (4;1, Ross, MacWhinney corpus)
 - f. Artoo and Three_Pee_O were soon picked up by Jawas (4;1, Ross, MacWhinney corpus)
 - g. The baby has to be carried by me (3;9, Emma, Weist corpus)

The examples in (47) show that children produce both passives with and without a *by*-phrase in the input. The examples provided are indeed those of intentional transitive verbs, as those are the verbs for which children have evidence of a productive rule in the input.

There also exists a line of work that entertains the hypothesis that the cross-linguistic variation in the acquisition of passives could result from differences in the input (e.g., Demuth 1989; Demuth, Moloi, and Machobane 2010). However, these analyses have yet

to consider the type frequency of the verbs in the input. Only the rate of passives based on the token frequencies of the verbs were considered. For instance, Maratsos et al. 1985 finds that full English passives are infrequent in the input. Kline and Demuth 2010 also suggest that the differences in the late acquisition of passives in English, and the early acquisition of passives in Sesotho may be the result of input differences. However, their investigation was based on a single corpus of English, the Brown Corpus (Brown 1973). Alcock, Rimba, and Newton 2012 also argue that the early acquisition of passives in two Eastern Bantu languages, Kiswahili and Kigiriyama, is due to a higher frequency of passives in the input. In this chapter, I provide a larger basis for this claim by examining all the corpus data available for North American English in CHILDES. Previous work has also attempted to attribute the asymmetry between actional and non-actional passives to the low frequency of non-actional passives in the input (Gordon and Chafetz 1990).

Along the same lines of the maturation hypothesis, Babyonyshev, Ganger, Pesetsky, and Wexler 2001 proposes that the late acquisition of passives can perhaps be attributed to the lack of external arguments, and not to the maturation of A-chains. In the previous chapters, I have illustrated how structures lacking an external argument are acquired. Raising constructions are acquired on the basis of non-referential expletive subjects, and as expletives are relative rare in the input, the raising property of these verbs emerges later on. In distinguishing unergative intransitives from unaccusative intransitives in chapter 4, we saw that children make use of notions such as intentionality to acquire this distinction. Moreover, there is no evidence that unaccusatives are acquired late. For instance, Vernice and Sorace 2018 show, using subject animacy, that children are aware of the distinction between unergative and unaccusative verbs around the age of 3. Although the property of having an external argument may indeed be readily available due to innate mechanisms, the claim cannot be made that children have difficulty with structures that lack an external argument. They may simply have to learn it from the input when their language shows evidence for it.

Finally, in regards to the passive asymmetry between actional and non-actional passives, one hypothesis is that there are semantic restrictions on passives (Maratsos et al. 1985; Pinker 1989; Pinker et al. 1987). Under this approach, mental verbs are claimed to be harder to passivize due to their semantic properties. One reason for the inability to passivize mental verbs is the Thematic Relations Hypothesis (Pinker et al. 1987). The Thematic Relations Hypothesis simply states that a verb can be passivized if it takes an object that is a patient. Therefore, verbs like *see*, which take a theme object, are not likely to be passivized by the learner. Exceptions within this class can be seen with verbs like *see* that can passivize. These types of exceptions are difficult to account for under the Thematic Relations Hypothesis, and one way of dealing with exceptions has been to allow for a more scalar restriction on passivization in these cases (Maratsos, Fox, Becker, and Chalkley 1985). A more scalar account proposes that verbs that take a theme as an object are simply less likely to passivize than verbs that take a patient object. A crucial component of this semantic restriction is that it is available to the learner from the start. This model of language learning also proposes that the verb semantics specify the verb's argument structure via linking rules. In this case, the semantic restriction on verbs results in a restriction on passivization (Pinker et al. 1987; Pinker 1989).

In this section, we have seen that there is some debate regarding when children acquire the passive construction; however, the evidence for children's knowledge of short actional passives is abundant (Hirsch and Wexler 2006, 2007; i.a.). Moreover, children have been shown to perform above chance on tests examining long passives of actional verbs with a *by*-phrase (e.g., Maratsos, Fox, Becker, and Chalkley 1985). In addition, the age of the acquisition of passives appears to vary cross-linguistically. Therefore, a theory of the acquisition of passives must be able to account for differences across languages. As pointed out by some (Bever 1970; Demuth 1989; i.a.), these differences can result from the nature of the linguistic input. Thus, in this chapter, I propose an analysis of the acquisition of passives that uses learning from positive evidence in the input. I also show how

the analysis accounts for the asymmetry between actional and non-actional passives.

4.3. Learning Passives from Positive Evidence

Previous approaches to the acquisition of passives have primarily argued for language-specific mechanisms that are innate to account for the developmental facts (Borer and Wexler 1987; Hirsch, Orfitelli, and Wexler 2007; Maratsos 1974; Pinker 1989; Pinker, Lebeaux, and Frost 1987). Moreover, as discussed in previous sections, the line of approach that argues for input-driven learning examines a small portion of the data without taking into consideration the number of verb types present in the input. In this section, I present an overview of an analysis that is driven by the number of verb types encountered by the learner.

The analysis for the acquisition of passives is as follows. The learner must first acquire verb classes in their language. Verb classes are formed according to the Active Mapping Model, which was described in Chapter 1. In the Active Mapping Model, the learner uses conceptual and structural cues to form verb classes. Two of the conceptual and structural cues that are immediately observable to the learner are that of intentionality and verb transitivity. In Chapter 3, the role of intentionality was outlined for the distinction between unergative and unaccusative verbs in English. The same kinds of verb classes come into play when considering the application of the passivization rule.

Infants as young as 7 months old have been shown to be attuned to intentionality at an early age (Carpenter, Akhtar, and Tomasello 1998; Woodward 1998). Hence, it is likely that intentionality plays a role in the acquisition of verb classes. Moreover, the valency of a verb is easily observable by the learner in the linguistic input, and therefore, can be used by the child while forming verb classes as well. Together, these cues form the following verb classes that are relevant to the acquisition of passive: intentional intransitives, intentional transitives, unintentional intransitives, and unintentional transitives.

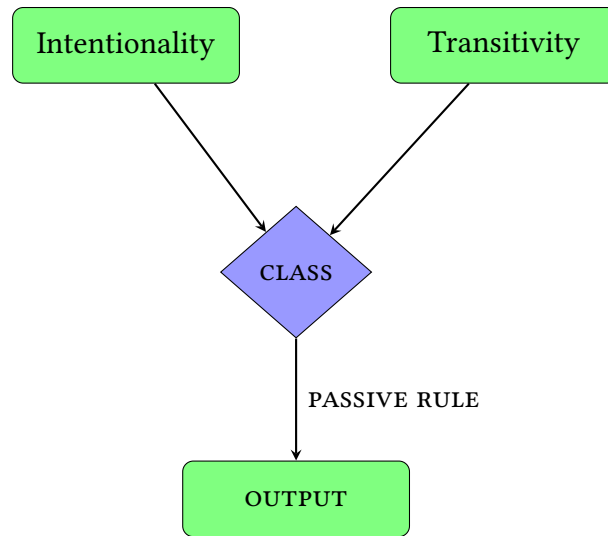


Figure 4.1: The acquisition of passives using the Active Mapping Model of acquisition. Conceptual and structural cues, in this case, intentionality and transitivity, both play a role in forming verb classes. The productivity of the passive rule is then evaluated at the level of the verb class.

The productivity of the passive rule within each verb subclass is then determined by the Sufficiency Principle (Yang 2016), as described in Chapter 1. Given the Sufficiency Principle, I argue that the passivization rule is only productive for the class of intentional transitives. The rule is not productive for the class of unintentional transitives, and therefore, the learner only discovers that a verb in this class is passivizable, if they encounter the verb used as a passive in the linguistic input. The unproductivity of the passivization rule for the unintentional transitive verb subclass is what results in what has been called the asymmetry between actional and non-actional verbs in children’s passives. In the following sections, I test these predictions.

4.4. Acquisition of Passives in English

In this chapter, I test two main aspects of the acquisition of passives in English: one, I investigate the age at which the passive rule becomes productive in English, given the input, and two, I examine whether the asymmetry in the learner’s knowledge between

actional and non-actional passives stems from the difference in the number of passives of each type in the input. To this end, I conduct two corpus-based studies. Study 1 examines the productivity of the verbal passives in the linguistic input, and Study 2 examines the learner's knowledge of the productivity of the passive. In both of these studies, I also analyze whether there is an asymmetry between actional and non-actional passives.

Following the analysis proposed in this dissertation, I argue that the developmental facts observed for the English passive fall out from the nature of the linguistic input. As a result, the actional and non-actional passive asymmetry is a consequence of learning from positive evidence in the input. The asymmetry between these two types of passives is not due to any restrictions on learning, or any semantic restrictions, but rather, due to the number of verbs of each type that undergo passivization. Here, I argue that the learner has verb classes such that the passivization rule is productive for the class corresponding to actional passives, but the rule is unproductive for the class of verbs corresponding to the non-actional passives. The passive rule is productive for a given class if and only if there is sufficient positive evidence in the input for the rule, as determined by the Sufficiency Principle.

Under the approach outlined above, the reason for why passives of non-actional verbs are learned late is because the passive rule for this verb class is unproductive in the input data. The number of passive verbs that occur in the input for the class of non-actional verbs are insufficient for the learner to have acquired a rule. Therefore, for the child to acquire a passive form for a verb in this class, the learner must hear each verb produced in a passive form in order to know that these verbs can be passivized. The number of actional verbs that passivize, however, are predicated to be sufficient enough for the learner to generalize a passive rule.

4.4.1. Study 1: Passives in Child-directed Input

4.4.1.1. Methods

When describing children's acquisition of passives, the literature often refers to verb classes as actional and non-actional (e.g., Borer and Wexler 1992). Others have described this division as actional verbs versus psychological (Pinker, Lebeaux, and Frost 1987) or mental verbs (Maratsos, Fox, Becker, and Chalkley 1985). The decision on what to call the verb classes becomes important when investigating each individual verb in corpus or empirical data. For instance, the verb *look after* does not refer to a physical act, and may be classified as non-actional by some, but it is not a psychological verb. When discussing this asymmetry in the acquisition of actional and non-actional passives, Maratsos et al. 1985 considered the following non-actional verbs: *see*, *hear*, *like*, *love*, *hate*, and *remember*. We see here that verbs like *see* and *hear* are not necessarily psychological, but they indeed are non-actional in the sense that they do not neatly correspond to actions in the physical world.

In this study, the verbs are classified as intentional or non-intentional. This distinction was used to differentiate between two types of intransitive verbs in previous chapters of the dissertation. When considering the verbs examined by Maratsos et al. 1985, *see*, *hear*, *like*, *love*, *hate*, and *remember*, we see that the conceptual property they have in common is that they lack intentionality. Children as young as 9 months are attuned to intentionality (e.g., Woodward 1998), and therefore, we can use this distinction to identify verb classes. By classifying verbs on the basis of intentionality, in this experiment, we test whether the developmental facts of the passive asymmetry can be accounted for under the proposed analysis.

Finally, there is a strong correlation between intentionality and transitivity, and verbs that are passivizable in English are known to correlate with transitivity (Nguyen and Pearl 2018). However, there are verbs that are intentional and not transitive, e.g., *jump*, *run*, and

therefore, both of these properties need to be included in our analysis of passives. Canonical transitive verbs passivize, and thus, in determining the properties that the learner might be attuned to, it is important to consider the transitivity of the verb. Hence, we take both intentionality and transitivity into account in our analysis of the acquisition of passives. The division of verb classes using transitivity and intentionality is illustrated in Figure 4.2.

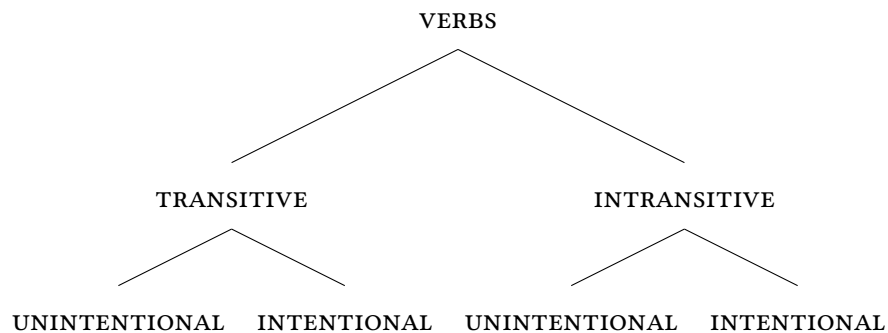


Figure 4.2: Classification of verbs using structural and conceptual cues.

To test the prediction of whether the asymmetry in the acquisition of actional versus non-actional passives stems from the input, I investigated the syntactic structures in which the 100 most frequent verbs in CHILDES occur. From the combined CHILDES input data, I extracted the 100 most frequent verbs. Verbs were tagged for part of speech and lemmatized using the Natural Language Toolkit. These verbs were first divided into verbs that occurred in a transitive frame and verbs that occurred only as intransitives. The transitive verbs were further divided into intentional and unintentional verbs. For each of these verbs, it was noted whether they occurred in a passive construction in the child-direct input. In order to determine whether the learner encounters a given verb as a causative, I examined these verbs in the combined input data available in all the corpora for North American English. Each caretaker’s speech in each file was used, resulting in a total of 6 million words of speech, which is about a year’s worth of input.

4.4.1.2. Results

The results show that the passivization rule in English is generalizable in the input for transitive verbs. Out of the 100 most frequent verbs, 22 of them occur only in an intransitive frame in the input. 13 of the 78 transitive verbs are unintentional, leaving 65 intentional transitive verbs. The summary of the number of verbs in each verb type is provided in Table 4.1.

Table 4.1: The 100 most frequent verbs in the CHILDES input data sorted by transitivity and intentionality.

Verb Type	Number of Verbs
Intentional Transitives	65
Unintentional Transitives	13
Intransitives	22

As shown in Table 4.2, out of the 22 intransitive verbs, none of them passivize. For the intentional transitives, 57 out of 65 passivize. For a class of 65 verbs, the Sufficiency Principle requires only 49 members to follow the rule ($N - N/\ln(N) = 65 - 16 = 49$); 57 verbs is well above that threshold. Turning to the unintentional transitives, we find that only 6 out of 13 of the verbs passivize. In this case, the Sufficiency Principle requires at least 8 verbs to follow the rule; 6 verbs out of 13 is below that threshold.

Table 4.2: The productivity of the passivization rule in child-directed input sorted by verb type. The threshold for each subclass is calculated using the Sufficiency Principle.

Verb Type	Number of Passives	Threshold	Rule Productive?
Intentional Transitives	57 out of 65	49	YES
Unintentional Transitives	6 out of 13	8	NO
Intransitives	0 out of 22	15	NO
Total	63 out of 100	78	NO

4.4.1.3. Discussion of the Results

The results indicate that the verbal passive is productive in the linguistic input. Moreover, when verbs are classified according to their structural and conceptual cues, which are transitivity and intentionality in this case, we see that the asymmetry between passives of verbs with different meanings is expected. The passive rule is productive for the class of intentional transitive verbs, but not for the class of unintentional transitives.

It should also be noted here that if the verbs are not categorized as shown above, the passive rule would not be productive in the input ($N - N/\ln N = 78$). As discussed in previous chapters, one of the reasons for the learner to seek subclasses may be due to the unproductivity of a rule for a larger class. The other reason for the learner to seek subclasses in this case could be simply because the learner is attuned to certain cues in the environment and the linguistic input, and can, therefore, use them in the language acquisition process.

The implications of the transitive verbs being categorized by intentionality are important. The verbs that are transitive, but unintentional, include *want*, *see*, *hear*, and *be*. Let us first discuss a verb like *be*, which is theoretically analyzed as a verb that does not introduce an external argument. The copula constructions in which *be* is observed are of the kind where there is a noun phrase preceding and following the word. A verb like *be* appears to be transitive, but it is in fact not a canonical transitive. Intentionality is one way in which a verb like *be* is distinguished from canonical transitives by the learner. On the other hand, subject experiencer verbs like *want* are analyzed as introducing an external argument; however, *want* resists passivization. One way in which this fact can be analyzed is by taking intentionality into account. If the passivization rule is not generalizable for unintentional transitives, then the child will only learn that verbs of this kind are passivizable if they hear the verb used in a passive construction. *Want* does not occur in a passive form in the input, and therefore, the learner does not assume that the verb passivizes.

Classifying verbs by intentionality also provides an analysis for the asymmetry observed in the acquisition of passives. The passivization rule is not productive for unintentional transitives, and therefore, the learner must hear the verb used in a passive form in order to know that an unintentional transitive verb passivizes. If the learner hasn't heard a verb of this type in a passive construction, they have no way of knowing that the verb can passivize, as the rule for this class is unproductive.

4.4.2. Study 2: Productivity of Passives in Adam's Grammar

Study 1 tested whether the English verbal passive construction is productive in the child-directed input data. In this second study, I test whether the learner's knowledge of the passivization rule in English is productive given the learner's vocabulary size. This section examines the productivity of the English verbal passive in Adam's grammar as a test case given Adam's vocabulary size. Adam from the Brown corpus was chosen for this study because a large amount of data is available for him. He also produces a number of short passives. Some examples of the passives that Adam produces are shown in (48).

- (48) a. Mommy (.) may I be (ex)cused [: excused]? (3;2)
b. So it can't be cleaned? (3;2)
c. It will be cooked in de [: the] minute (3;3)
d. Dat [: that] dolly was made in Italy? (3;7)
e. Saw the cows being milked (3;10)

Since Adam produces a number of passives, his data make the perfect test case to examine whether the construction is productive in his grammar.

4.4.2.1. Methods

In order to test whether Adam has a productive passive rule, the point when he produced his first verbal passive, which was at age 3;2, was identified. Then, to examine whether

the passive rule was productive for Adam at that time, I first estimated his vocabulary size by counting each verb produced by Adam up to age 3;2. This resulted in a total of 208 verbs, which were divided according to transitivity and intentionality. I then checked to see whether each verb occurred in a passive construction in the input. The input data used were the speech of the caregivers in the combined corpora of North American English. This resulted in a total of roughly 6 million words of speech. This method was chosen over only examining Adam's input data to make up for the small sample size. It is unlikely that a verb occurs in all possible constructions in the input data for a single child, but by examining the combined input data for child-directed speech, we can get a better sense of what a learner is likely to hear in the early stages of the acquisition period.

4.4.2.2. Results

Adam's verbs were categorized as intentional transitives, unintentional intransitives, and intransitives. Out of 145 intentional transitives produced by Adam, 117 of the verbs occurred as a passive in the child-directed input. The results indicate that Adam has a productive verbal passivization rule in his grammar, as the Sufficiency Principle only requires 116 ($N - \ln(N) = 145 - 29 = 116$). In contrast, only 5 out of the 14 unintentional transitives, and 0 out of the 49 intransitives occurred in a passive construction. For a class of 14 unintentional transitives, the Sufficiency Principle requires 9 members to follow the rule in order for the rule to be considered productive ($N - \ln(N) = 14 - 5 = 9$). A summary of Adam's verbs is provided in Table 4.3.

Table 4.3: Adam's verb types and the productivity of the passive rule. The threshold for each subclass is calculated using the Sufficiency Principle.

Verb Type	Number of Passives	Threshold	Rule Productive?
Intentional Transitives	117 out of 145	116	YES
Unintentional Transitives	5 out of 14	9	NO
Intransitives	0 out of 49	36	NO

4.4.2.3. Discussion of the Results

The results show that given Adam's vocabulary size, he has a productive verbal passive rule for the class of intentional transitive verbs, but not for the class of unintentional transitives. For a class of 145 verbs, the Sufficiency Principle requires that at least 116 verbs follow the rule. In this case, 117 of the 145 intentional transitives were observed to follow the passivization rule. This result, thus, appears to be barely over the threshold of productivity. However, if the verbs are arranged by frequency, we see that the rule is well over the sufficiency threshold; it only appears to be less productive for the slightly less frequent verbs. The number of passives seen in the input with Adam's verbs sorted by frequency are shown in Figure 4.3. In Figure 4.3, the number of verbs needed to follow the rule along with the number of verbs that follow the rule in Adam's vocabulary are illustrated.

As seen in Figure 4.3, there is substantial evidence in the input for the passive rule for the relatively more frequent verbs in CHILDES. For the 10 most frequent verbs that Adam knows, 9 of them occur as a passive in the input when the Sufficiency Principle only requires 6. Similarly, 19 out of 20 of the most frequent verbs occur as a passive when only 14 are required. Even for the 100 most frequent verbs, 86 of them occur in a passive frame in the input, which is well above the Sufficiency Principle's threshold of 78. Thus, we see that based on the input data, children have sufficient positive evidence to form the passive rule for the class of intentional transitives.

The results obtained here indicate that the verbal passive is productive early on for a child learning English. We find that the passive rule for Adam is productive by age 3;2, which is around the time the passive construction has been observed to be learned by Sesotho-speaking children. Often in the literature, we see that these languages are cited as having a different developmental trajectory. The results presented from this corpus study indicate that this might not be the case.

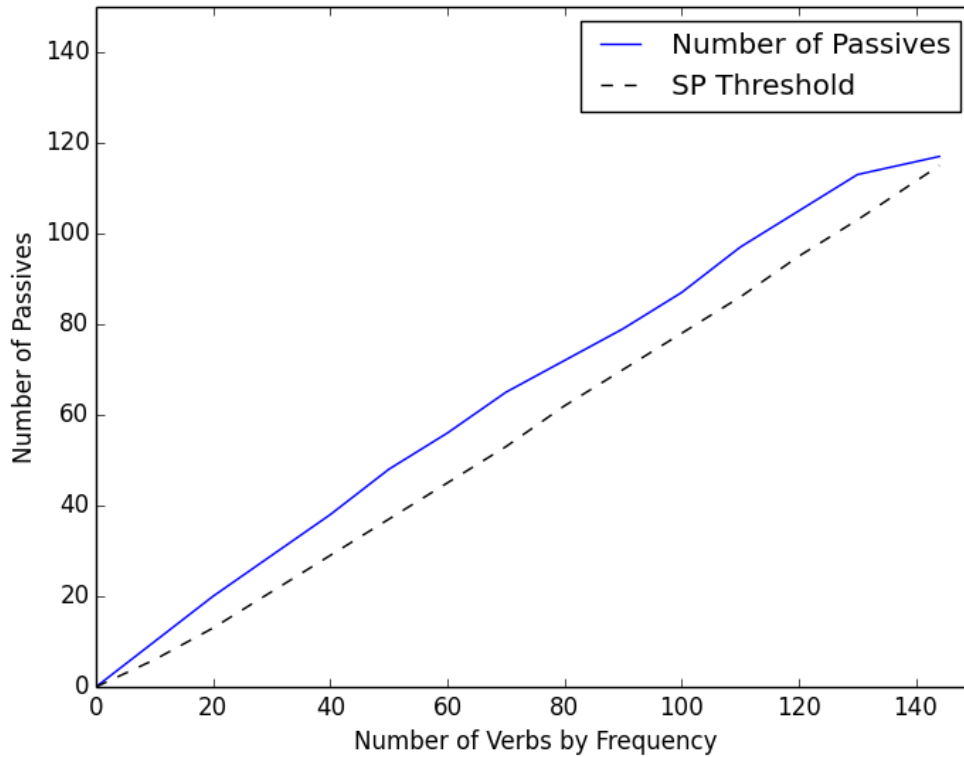


Figure 4.3: Adam’s knowledge of passives.

4.5. Acquisition of Passives in Sesotho

The acquisition of passives in Sesotho gained attention in light of the maturation hypothesis. The early acquisition of Sesotho is a known and established counterexample to the claim that A-chains mature late. In this section, I examine whether the proposed analysis captures the acquisition facts for Sesotho-learning children. Although the passive asymmetry has not explicitly been identified for Sesotho, we build on our analysis for English in order to investigate passives in Sesotho. For English learning children, we have seen that the asymmetric acquisition of passives stems from the lack of passives of non-actional verbs in the input. Here, I examine this pattern in Sesotho as well.

4.5.1. Background

The early acquisition of Sesotho passives was first shown in Demuth 1989. In Demuth 1989, several instances of full passives that included a *by*-phrase in the utterances of three children are provided. Some of these examples are illustrated in (49).

- (49) a. se-khann-o-a ke 'na
SM-drive-PASS-M by me
'It is driven by me.' (2;6, Keneuoe, Demuth 1989:73)
- b. 'Na ke-kut-uo-e ke nkhono oaka.
PN SM-cut.hair/PRF-PASS-M by grandmother my
'As for me, I've been given a hair-cut by my grandmother.' (2;8, Keneuoe, Demuth 1989:62)
- c. ke-tla-be ke-tlo-rut-o-a ke 'Me Mamojela
SM-FUT-be SM-FUT-teach-PASS-M by Mrs. M
'I'm going to be taught by Mrs. M.' (3;2, Litlhare, Demuth 1989:73)

Further work on Sesotho passives (Kline and Demuth 2010) also reveals that a large number of verb types were used as passives by both children and adults. Kline and Demuth 2010 find that out of the 429 verb types used by adults, 74 were used in a passive construction. In the child production data, 90 of 525 verb types occurred in a passive frame. This data indicates that Sesotho learning children indeed have a good mastery of the passive construction.

In addition to being able to produce and comprehend passives, Sesotho learning children also have the passive construction as a productive rule in their grammar. The productive passive rule is evident from the learner's ability to extend the passive rule to novel verbs. Demuth, Moloi, and Machobane (2010) tested 3 year old Sesotho learning children's ability to generalize the passive construction to novel verbs. They found that children were able to generalize the passive construction to novel verbs with the semantics of actional verbs.

Although passives were ubiquitous in both the adult and child production data, Kline and Demuth 2010 find that only 4% of the adult passives and only 5% of children's passives were those of non-actional verbs. In contrast, the percentage of non-actional verbs in active constructions is much larger. 22% of the verbs for adults, and 24% of the verbs for children, were non-actional in the active construction. However, the rate of passives in the input is not informative regarding the productivity of the construction. Thus, the productivity of non-actional passives for Sesotho learning children remains an open question, one which is addressed in the following section.

4.5.2. Sesotho-Learning Children's Knowledge of Passives

Even though there is an abundance of evidence for Sesotho learning children's knowledge of actional passives, it is unclear whether there is an asymmetry between passives of actional and non-actional verbs. More specifically, even though children produce passives of non-actional verbs, is the passivization rule for these verbs productive in their grammar? Evidence from experimental and corpus data illustrate that children can produce and comprehend passives of non-actional verbs that they have previously encountered in the input (Demuth, Moloji, and Machobane 2010). In this section, we investigate whether the input allows for the passivization rule to be generalizable to novel non-actional passive verbs. This distinction is crucial in understanding whether Sesotho learning children are only able to produce passive forms of non-actional verbs that they have previously encountered, or whether they have learned a productive rule. Addressing this question also supports our conclusion for English learning children, where we found that the input data is such that passives of non-actional verbs have to be learned on an item-by-item basis.

4.5.2.1. Methods

To investigate whether the passive rule is productive in the input for children learning Sesotho, each caregiver's speech for all of the 4 children in the CHILDES database for Sesotho were examined. From this input data, each verb type produced by the parent was extracted from the corpus using a python script. The verbs were extracted based on the translation of the verb root provided in the morphological coding. To ensure that translation of the root of the verb extracted was accurate, I also checked each of these verbs by hand with the original Sesotho text using the provided glosses.

To determine the most frequent verbs in the child-directed input the frequency for each verb was obtained. These verbs were then coded for intentionality and transitivity. The transitivity of the verb was determined from its use in the input to avoid assuming the same argument structure for the Sesotho verbs as English. Each verb obtained was then searched for in the input data. A verb was noted to have occurred as a passives if it occurred as a passive in the caregiver's speech at least once.

4.5.2.2. Results

The results indicate that there is insufficient evidence in the input for the child to assume that the rule is productive. Out of the 100 most frequent verbs in child-directed speech, 78 were intentional transitives, 8 were unintentional transitives, 6 unintentional intransitives, and 8 were intentional intransitives. Out of 78 intentional transitives, only 54 of them occurred in a passive construction in the input. 3 out of 8 of the intentional intransitives occurred as a passive. None of the 8 unintentional transitives, and none of the 6 unintentional intransitives occurred as a passive in the input data. These results are summarized in Table 4.4.

Table 4.4: Passives by verb type in the Sesotho child-direct input. The threshold of generalization is calculated using the Sufficiency Principle.

Verb Type	Number of Passives	Threshold	Rule Productive?
Intentional Transitives	54 out of 78	60	NO
Intentional Intransitives	3 out of 8	4	NO
Unintentional Transitives	0 out of 8	4	NO
Unintentional Intransitives	0 out of 6	3	NO

4.5.2.3. Discussion of Sesotho Passives

The results of the corpus data analysis do not appear to support the productivity of the Sesotho passive. For each of the four verb subclasses, the number of verbs that occur in a passive construction is below the threshold of the Sufficiency Principle. These findings appear to be in contrast to previous results (e.g., Demuth et al. 2010), which suggest that Sesotho-learning children productively use the passive rule, at least for actional verbs.

Although the number of verbs that occur as passive are well below the threshold of the Sufficient Principle, it is likely that the size of the corpus examined plays a role in the results obtained. Many of the verbs examined occurred only a few times in the input, and therefore, it is unsurprising that they were not used as a passive. To illustrate this point, the size of the input data corpus examined for English contained about 6 million words of speech. In contrast, the Sesotho input data contains only 101,502 Sesotho words.

By examining the verbs by frequency, we can see whether a verb could have occurred in a passive frame had it occurred more times in the input. Kline and Demuth 2010 claim that passives of actional verbs are more frequent; therefore, as a test case, we can examine the class of intentional verbs here, which has a large overlap with Kline and Demuth 2010's class of actional verbs. To determine whether there is a correlation with frequency in the productivity of the passive, I examined the 100 most frequent intentional verbs in the input. The results of this test case are illustrated in Figure 4.4.

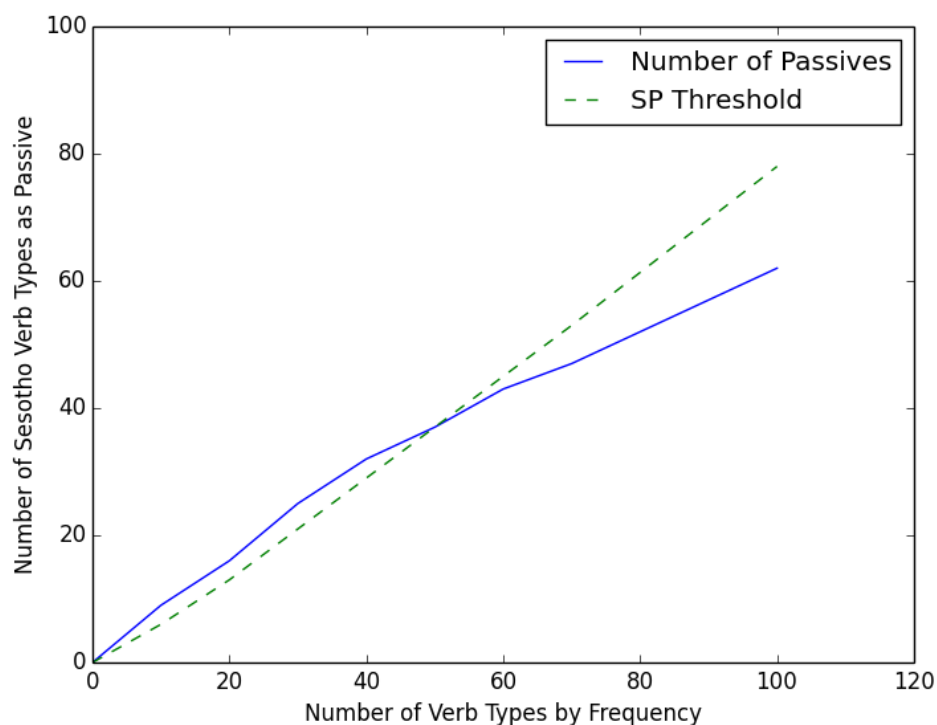


Figure 4.4: Number of passives in the input arranged by frequency.

Figure 4.4 illustrates the 100 most frequent verbs in the input sorted by frequency. At each interval, the number of verbs that occurred as a passive in the input data can be seen. In this graph, we see that the passive rule is productive in the input for the first 50 verbs. After this point, the rate of passives used decreases. This result is consistent with the idea that if the less frequent verbs occurred more times in the input, they would have been used in a passive construction as well.

The results obtained from the corpus study are not in line with the predictions of the proposal put forth in this chapter. However, it has been long established that Sesotho-learning children master passives by age 3. Thus, even though the results are not in line with the predictions of the analysis, it is likely because there was not enough corpus data available. In comparing Sesotho to English, Sesotho is the established baseline for the early acquisition of the passive construction.

4.6. Discussion

Previous research on the acquisition of passives claims that children learning Sesotho acquire the passive construction earlier than English-learning children. In the discussion of the acquisition of passives in English, it has also been observed that children learn passives of actional verbs before passives of non-actional verbs. These previous observations left open the question of whether children learning English are at a disadvantage when acquiring passives compared to Sesotho-learning children.

The results presented in this chapter suggest that perhaps children learning both Sesotho and English acquire actional passives before non-actional passives. Under the analysis presented in this dissertation, the observed distinction in the verb classes is in fact between intentional and unintentional verbs. The results suggest that the difference between the two languages is due to the observation that the passive construction is not a generalizable rule for unintentional verbs; each verb belonging to this sub-type must be learned individually from its use in the input.

In this chapter, Study 1 first investigated verbal passives in the 100 most frequent verbs in the child-directed input. The results from this study showed that the verbal passive is completely productive in the input data. Furthermore, the results of Study 2, which tested the productivity of the passive rule in Adam's grammar, indicate that the passive rule is generalizable for intentional transitive verbs by the age of 3. The findings of these studies indicate that the arguments for children acquiring English passives after the age of 5 are not tenable.

The claim that English passives are acquired late also stemmed from theoretical considerations; i.e., the maturation hypothesis (Borer and Wexler 1987). The maturation hypothesis, which claims that A-chains do not mature until age 5, precludes early child passives. However, the data presented here, along with previous empirical work on children's passives (e.g., Pinker, Lebeaux, and Frost 1987), indicates that the maturation hypothesis

cannot be correct. The examples in (47) also show that the claim that English-speaking children cannot produce passives before the age of 5 is based on insufficient data.

The studies presented in this chapter also address the asymmetry in children's knowledge of passives (Maratsos et al. 1987, Pinker et al. 1987; i.a.). In this chapter, I have argued that children use both structural and conceptual cues in forming verb classes. Children then observe the passive rule for each subclass in the input. In this case, children learn that the passive rule is productive for the class of intentional transitive verbs, but not unintentional intransitive verbs. These two classes characterize the asymmetry observed in the learner's ability to passivize verbs like *kick*, but not verbs like *see*.

At this point, we see that learning the passive construction is much like learning the causative alternation rule, which was discussed in the previous chapter. It should be noted here that the learning trajectory of the two constructions feed into each other to a certain degree. In Chapter 3, I showed that children overgeneralize to use unintentional intransitive verbs as transitives in the input. For instance, the learner may use the unintentional intransitive verb *fall* as a transitive. This transitive use of *fall* may be intentional. In that case, the learner may include *fall* in their calculations of intentional transitive verbs, and thus, the learner may incorrectly passivize a verb like *fall*. This may account for the few overgeneralizations of passives that have been observed (e.g., Pinker 1989), in particular with the verb *die*. This overgeneralization vanishes, however, once the learner retreats from their overgeneralization of the causative alternation rule.

In this chapter, the overgeneralization of the passive form has not been discussed as there are not many instances of overgeneralization seen in the child production data. Nevertheless it may be possible that there is an occasional overgeneralization of the passive form. This overgeneralization stems from positing an intentional transitive form of the verb, which results from the overgeneralization of the causative alternation discussed in Chapter 3. There is no independent overgeneralization of the passive construction, as all intentional transitive verbs can passivize, and verbs that can passivize in other classes

must be learned individually from the input.

There is one final point of discussion to be made in regards to Sesotho passives versus English passives. In Sesotho, there is a distinct passive morpheme which indicates the existence of a separate construction. In English, the auxiliary combined with the past participle denotes that meaning. It may be possible that there is some effect of the distinct passive morphology in Sesotho on the learning process, but English-learning children appear to be aware of the passive construction in their language. This is seen, for instance, in the overgeneralization of the past participle in Adam's passives.

- (50) a. I want to be shot [: shot] [* +ed] (3;8)
b. How could it go up if it's not (.) if it's not flied [: flew] (3;10)

As seen in (50), Adam makes morphological errors with the past participle form, but uses it in the passive construction regardless. The overgeneralization of the morphology suggests that children learning English are aware of the passive construction even though it uses the past participle form. Hence, here we assume that the morphology does not play a large role in hindering the learning process in English. Thus, any potential effect of the morphology between the two languages is not discussed any further here.

4.7. Conclusion

The findings presented in this chapter also indicate that children learn the passive construction from positive evidence in the input. They use structural and conceptual cues in the input to form verb classes. The rules observed in the input for a given verb class are then generalized if there is sufficient evidence in the input. If there is insufficient evidence in the input for a rule, the learner must learn the properties of the class on an item-by-item basis.

The analysis discussed in this chapter, and throughout the dissertation, does not presuppose that the learner is inherently aware of the semantic restrictions on each verb.

Moreover, a model where the semantics of a verb are inherently linked to its argument structure is not required in order to account for the patterns observed in the acquisition of the passive construction. In this chapter, I have shown that the passive construction is productive in English, despite previous claims in the literature (e.g., Hirsch and Wexler 2006). I have also shown that the asymmetry between passives of intentional and unintentional verbs is expected given the number of verbs of each type that passivize. Furthermore, despite the analyses of the early acquisition of Sesotho found in the literature, the Sesotho corpus study conducted here did not support productivity of the passive rule in the input. However, a larger corpus is needed to successfully analyze Sesotho passives in the input.

Chapter 5

Conclusions

In this dissertation, I examined the acquisition of verb argument structure in three cases: the acquisition of raising and control verbs, the acquisition of the causative alternation, and the acquisition of passives. I used two learning models, the Sufficiency Principle (Yang 2016) and the Active Mapping Model of Language Acquisition proposed in Chapter 1, to analyze the aforementioned case studies. In the subsection below, I summarize the findings of the previous chapters, and then conclude with some final remarks.

5.1. Summary

Chapter 2 showed that it is possible to mistreat the problem of overgeneralization if the learner is assumed to generalize from insufficient data. Sufficient evidence was determined via the Sufficiency Principle. Using the case of raising, control, and ambiguous verbs, I illustrated that the acquisition of these verbs does not lead to a learnability problem, from which the learner eventually needs to retreat. The learner does not assume that because some predicates can be both raising and control, that all potential raising and control predicates can be both. By examining the number of control, raising, and ambiguous verbs in the input data, we see that the learner would not overgeneralize the class of ambiguous verbs to assume that all verbs can be either raising or control. Instead, the learner distinguishes between these classes using positive evidence from the input, which in this case is in the form of non-referential subjects.

Chapter 3 examined the problem of overgeneralization in the case of the acquisition of causatives. Here, we saw evidence from production data that children overgeneralize the causative alternation (e.g., *the vase broke/John broke the vase*) to verbs that are pure intransitives in the adult grammar (e.g., *John fell/*I fell John*). In the acquisition of causatives, I demonstrate that the overgeneralization errors are predicted under the Active Mapping Model where the learner categorizes verbs into classes based on conceptual and structural cues. Given the learner's vocabulary size and verb classes, the causative alternation rule is found to be productive when the input the learner receives is examined. Thus, under this learning model, the child overgeneralization errors are predicted. Moreover, using the Sufficiency Principle to determine the threshold of generalization, I showed that the learner retreats from overgeneralization when their vocabulary size increases, as the rule is then no longer productive when the input is examined.

In Chapter 3, the Sufficiency Principle was also tested experimentally. 45 adults and 26 children were tested using an artificial language paradigm over three experiments to determine their threshold of generalization. Participants were assigned to one of two language groups, one where 8 verbs out of 10 occurred as transitive causatives, and two verbs as pure intransitives, and the other where 5 out of 10 verbs occurred as transitive causatives, and five verbs as pure intransitives. The Sufficiency Principle predicts overgeneralization for the 8 out of 10 condition, but not the 5 out of 10 condition. The verbs in the experiments were presented using a Zipfian distribution. This variation in frequency allowed for testing of an entrenchment-based learning approach. The results showed that adults produced the transitive form consistently for each verb. In contrast, child participants overgeneralized significantly more in the 8 out of 10 condition, as opposed to the 5 out of 10 condition. These results are in line with the prediction of the Sufficiency Principle. Furthermore, no frequency effects were found, suggesting evidence against an entrenchment-based learning approach.

Finally, in Chapter 4, I used the Sufficiency Principle and the Active Mapping Model

to examine the developmental trajectory of children's acquisition of passives. This chapter sought to answer two questions: whether the passive construction is productive for the learner early on, and whether the asymmetry in the acquisition of actional and non-actional passives (e.g., Pinker et al. 1987) can be accounted for under the models of language learning assumed in this dissertation. Using the child production data and the input data, I showed that the passive construction is productive in the input, and productive for the English-learning child. By examining Adam's vocabulary and verb classes, the passive rule was found to be productive for Adam given the number of passives in the input. Using the Active Mapping Model, I showed that the asymmetry in the acquisition of passives is predicted.

5.2. Final Remarks

This dissertation has covered three main themes throughout. The first theme discussed is how children come to acquire verb argument structure. In regards to this question, I addressed two main theories of language learning, a frequency-based indirect negative evidence approach, and a learning from positive evidence-based approach. I showed that an indirect negative evidence approach is ineffective because it requires a large amount of data in the input that is unavailable to the learner. Additionally, I provided child data that indicates that the learner does not follow the developmental trajectory proposed by an indirect negative evidence account. This theme brings us to the second aspect of language acquisition discussed in the literature: the problem of overgeneralization.

For the problem of overgeneralization, when the learner has acquired a grammar that is a superset of the adult grammar, the question is how can the learner retreat from the superset grammar in the absence of direct negative evidence. In relation to the problem of overgeneralization, two lines of approaches have been proposed. Either the learner is said to retreat using indirect negative evidence (e.g., Ambridge 2008), or retreat using positive

evidence (Yang 2016). This dissertation supports the second line of approach.

Finally, the third theme discussed throughout the dissertation is how the learning models proposed in the literature fair when accounting for some of the hardest problems of language acquisition in the domain of verb learning. One of the theories, the indirect negative evidence approach, was already mentioned above. The other approach is Pinker 1989's innate linking rules approach. Under the linking rules approach, the learner is born with innate mappings between the semantics and the syntactic structure. Upon learning the verb semantics, the learner forms verb classes and is readily equipped with a syntactic structure associated with that class of verbs. One indispensable problem with the innate linking rules account is that many of the overgeneralization errors observed are not predicted in this learning model. Therefore, in this dissertation, I propose a new model of language learning, the Active Mapping Model, where the learner forms classes based on independent conceptual and structural cues. Any mapping that is formed between the syntax and semantics is acquired through the learner's linguistic experience. Using this model, I showed that the learner's developmental trajectory can be predicted; i.e., the Active Mapping Learning Model is compatible with the data observed for the child's acquisition of verb argument structure.

Given the findings of this dissertation, there is one important area of work that still needs to be pursued in the domain of the acquisition of verb classes. In the previous chapters, I described some potential structural and conceptual cues that can be used by the learner to form verb classes, but the nature of these cues must be explicated in future work. Crucially, the cues the learner uses to form classes must be distinct from the structural rules observed in the input that apply to the class. The range of potential conceptual and structural cues has not been addressed in its entirety here.

A final point that should be made in regards to the Active Mapping Model is the way in which it can be tested to be proved accurate or inaccurate. Currently, I have shown that the Active Mapping Model best accounts for the child developmental facts outlined in the

previous chapters. However, here, I would like to lay out one potential way in which the Active Mapping Model is falsifiable. If the developmental data found is different from what is predicted using the conceptual and structural cues available in the input, then that would constitute as evidence against a primarily input-driven model of language learning. In other words, if we find that the learner's linguistic behavior is different from what we would expect given the input data, then we may have evidence for an internal factor that drives the learner's process of language acquisition.

To conclude, the Active Mapping Model is proposed here to begin asking questions about language acquisition from a different perspective. Can the learner's developmental trajectory be accounted for solely from the input data, using a model that is informed by linguistic theory? This dissertation is a step in that direction.

Appendix A

Chapter 2 Appendix

Subject control verbs with total number of occurrences in the North American English CHILDES input data (42 total):

bother (351), claim (5), come (23575), leave (4119), look forward to (12), neglect (5), petition (4), pretend (1698), wait (5844), care (783), like (33845), prepare (44), pretend (1693), scramble (82), see (56412), think (22628), volunteer (10), want (32363), attach (190), decide (376), forget (1572), hope (798), plan (174), stick (496), try (6323), use (4584), agree (78), attempt (5), hate (348), offer (64), pay (452), refuse (13), arrange (26), choose (220), learn (920), long (7), mean (5360), seek (7), dare (43), desire (5), ask (2724), vow (2)

Raising verbs with total number of occurrences in the North American English CHILDES input data (14 total)

appear (58), get (45079), happen (5567), proceed (3), seem (464), start out (17), stay (3009), tend (39), about (148), likely (13), sure (35), going (13552), set (1133), bound (8), suppose (1659)

Ambiguous control and raising verbs with total number of occurrences in the North American English CHILDES input data (12 total):

fail (14), begin (468), continue (463), manage (38), need (8714), promise (200), stop (3429), grow (809), start (2337), threaten (12), have (47257)

Appendix B

Chapter 3 Appendix

Adam's Verb List (208 total):

go, come, laugh, cry, fall, change, shake, move, break, stay, hurt, turn, drop, pop, open, fit, grow, tickle, live, blow, shine, push, drip, lose, tear, roll, twinkle, ring, close, bake, fold, taste, hang, call, slip, work, crack, burn, bump, feel, bend, wind, flash, tip, hide, smash, sail, blast, print, dry, sneeze, play, sit, stand, jump, lie, walk, drive, climb, camping, swim, sweep, bite, work, crawl, wait, listen, hurry, run, dance, eat, sleep, squeak, point, paint, fly, sing, squeeze, wake, fight, kiss, talk, finish, pack, squeal, dig, care, growl, tease, pay, exercise, stir, swing, pinch, quack, hug, swallow, rest, leave, rock, shave, back, fire, smoke, dress, scream, marry, sow, ski, fish, look, put, read, get, like, hop, see, remember, hit, be, find, give, beat, ride, write, wipe, comb, want, drink, draw, pull, pick, fix, take, throw, kick, watch, stop, buy, check, wear, keep, happen, matter, scratch, cook, step, show, knock, catch, use, pour, carry, need, build, dip, fasten, save, hold, make, do, have, miss, bring, help, press, hand, chew, park, know, shoot, ask, let, cut, try, say, spank, count, excuse, reach, frighten, dump, stuck, skip, pretend, think, mix, hate, pat, tie, tell, rope, mail, spell, learn, lift, cross, wash, zip, seem, mock, match, rustle, punch, screw, plug, serve, spill

Ross's Verb List (121 total):

go, come, cry, turn, move, drop, shut, stay, close, fall, hurt, live, open, feel, call, sit, break, twist, rip, slip, tickle, tumble, work, pop, do, play, try, nap, see, kiss, look, finish, drive, read, jump, yell, leave, sleep, fool around, dare, bite, eat, scratch, work, stand, blow, spit,

fly, swim, burp, talk, growl, sing, run, jog, ride, wake, fight, climb, whisper, point, knock, smile, wait, scream, forget, know, be, fix, like, give, get, have, take, put, think, hurry, click, want, wear, let, say, need, show, help, carry, buy, bug, hit, find, cut, feed, pull, hold, touch, make, throw, record, push, kill, drink, steal, flush, scare, hear, transform, repeat, taste, salt, stop, shoot, pick, happen, bump, use, brush, spank, love, list, cook, bring

Appendix C

Chapter 4 Appendix

100 most frequent verbs in the CHILDES input data:

do, put, say, make, take, eat, tell, give, play, find, read, call, turn, help, hold, show, use, leave, play, keep, open, finish, wear, hurt, throw, buy, write, ask, break, watch, stop, fix, move, cut, sing, draw, pull, hit, bring, push, build, fit, wash, run, start, pick, catch, touch, set, stick, clean, cause, dress, bite, close, drive, scare, get, try, drink, work, feel, walk, learn, ride, go, come, think, look, let, sit, happen, talk, fall, wait, sleep, stay, stand, suppose, guess, bet, live, pretend, wonder, hope, cry, listen, see, know, like, hear, love, understand, be, have, need, mean, remember, forget, want

Bibliography

- Alcock, Katherine J, Kenneth Rimba, and Charles RJC Newton (2012). “Early production of the passive in two Eastern Bantu languages”. In: *First Language* 32.4, pp. 459–478.
- Alishahi, Afra and Suzanne Stevenson (2008). “A computational model of early argument structure acquisition”. In: *Cognitive science* 32.5, pp. 789–834.
- Allen, Shanley EM and Martha B Crago (1996). “Early passive acquisition in Inuktitut”. In: *Journal of Child Language* 23.1, pp. 129–155.
- Ambridge, Ben (2013). “How Do Children Restrict Their Linguistic Generalizations? An (Un-) Grammaticality Judgment Study”. In: *Cognitive Science* 37.3, pp. 508–543.
- Ambridge, Ben, Amy Bidgood, Julian M Pine, Caroline F Rowland, and Daniel Freudenthal (2016). “Is passive syntax semantically constrained? Evidence from adult grammaticality judgment and comprehension studies”. In: *Cognitive science* 40.6, pp. 1435–1459.
- Ambridge, Ben and Elena VM Lieven (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press.
- Ambridge, Ben, Julian M Pine, Caroline F Rowland, Rebecca L Jones, and Victoria Clark (2009). “A Semantics-Based Approach to the “No Negative Evidence” Problem”. In: *Cognitive Science* 33.7, pp. 1301–1316.
- Ambridge, Ben, Julian M Pine, Caroline F Rowland, and Chris R Young (2008). “The effect of verb semantic class and verb frequency (entrenchment) on children’s and adults’ graded judgements of argument-structure overgeneralization errors”. In: *Cognition* 106.1, pp. 87–129.

- Ammon, Mary Sue Hersey (1980). "Development in the Linguistic Expression of Causal Relations: Comprehension of Features of Lexical and Periphrastic Causatives." PhD thesis. University of California, Berkeley.
- Aslin, Richard N and Elissa L Newport (2012). "Statistical learning: From acquiring specific items to forming general rules". In: *Current directions in psychological science* 21.3, pp. 170–176.
- Austin, Alison (2010). "When children learn more than what they are taught: regularization in child and adult learners". PhD thesis. University of Rochester.
- Babyonyshev, Maria, Jennifer Ganger, David Pesetsky, and Kenneth Wexler (2001). "The maturation of grammatical principles: Evidence from Russian unaccusatives". In: *Linguistic inquiry* 32.1, pp. 1–44.
- Baker, Carl L (1979). "Syntactic theory and the projection problem". In: *Linguistic Inquiry* 10.4, pp. 533–581.
- Baron, Naomi S (1972). *The evolution of English periphrastic causatives: contributions to a general theory of linguistic variation and change*. Committee on Linguistics.
- Bates, Elizabeth and Brian MacWhinney (1987). "Competition, variation, and language learning". In: *Mechanisms of language acquisition*, pp. 157–193.
- Becker, Misha (2006). "There began to be a learnability puzzle". In: *Linguistic Inquiry* 37.3, pp. 441–456.
- (2009). "The role of NP animacy and expletives in verb learning". In: *Language Acquisition* 16.4, pp. 283–296.
- (2014). *The acquisition of syntactic structure: Animacy and thematic alignment*. Vol. 141. Cambridge University Press.
- Berwick, Robert C (1985). *The acquisition of syntactic knowledge*. MIT press.
- Bever, Thomas G (1970). "The cognitive basis for linguistic structures". In: *Cognition and the development of language* 279.362, pp. 1–61.

- Boersma, Paul (2002). "Praat, a system for doing phonetics by computer". In: *Glot international* 5.
- Boguraev, Bran and Ted Briscoe (1987). "Large lexicons for natural language processing: utilising the grammar coding system of LDOCE". In: *Computational Linguistics* 13.3-4, pp. 203–218.
- Borer, Hagit (1994). "The projection of arguments". In: *University of Massachusetts occasional papers in linguistics* 17, pp. 19–47.
- Borer, Hagit and Kenneth Wexler (1987). "The maturation of syntax". In: *Parameter setting*. Springer, pp. 123–172.
- (1992). "Bi-unique relations and the maturation of grammatical principles". In: *Natural Language & Linguistic Theory* 10.2, pp. 147–189.
- Bowerman, Melissa (1982). "Starting to talk worse: Clues to language acquisition from children's late speech errors". In: *U shaped behavioral growth*. Academic Press, pp. 101–145.
- (1988). "The 'no negative evidence' problem: How do children avoid constructing an overly general grammar?" In: *Explaining language universals*. Basil Blackwell, pp. 73–101.
- Bowerman, Melissa and William Croft (2008). "The acquisition of the English causative alternation". In: *Crosslinguistic perspectives on argument structure: Implications for learnability*, pp. 279–307.
- Braine, Martin DS and Patricia J Brooks (1995). "Verb argument structure and the problem of avoiding an overgeneral grammar". In: *Beyond names for things: Young children's acquisition of verbs*. Lawrence Erlbaum Associates, Inc, pp. 353–376.
- Braine, Martin DS et al. (1990). "Exploring language acquisition in children with a miniature artificial language: Effects of item and pattern frequency, arbitrary subclasses, and correction". In: *Journal of memory and language* 29.5, pp. 591–610.

- Brooks, Patricia J and Michael Tomasello (1999). “How children constrain their argument structure constructions”. In: *Language*, pp. 720–738.
- Brown, Roger (1973). *First language: The early stages*. Harvard U. Press.
- Cairns, Helen Smith, Dana McDaniel, Jennifer Ryan Hsu, and Michelle Rapp (1994). “A longitudinal study of principles of control and pronominal reference in child English”. In: *Language*, pp. 260–288.
- Carey, Susan and Elizabeth Spelke (1994). “Domain-specific knowledge and conceptual change”. In: *Mapping the mind: Domain specificity in cognition and culture*, pp. 169–200.
- Carpenter, Malinda, Nameera Akhtar, and Michael Tomasello (1998). “Fourteen-through 18-month-old infants differentially imitate intentional and accidental actions”. In: *Infant behavior and development* 21.2, pp. 315–330.
- Chen, Zhuo, Virginia Valian, and M. Chodrow (2016). “Subject Production by Monolingual English-speaking Children: A Corpus Study”. Poster at Boston University Conference on Language Development.
- Chomsky, Carol (1969). *The acquisition of syntax in children from 5 to 10*. MIT Press.
- Clark, Eve V (1978). “Awareness of language: Some evidence from what children say and do”. In: *The child’s conception of language*. Springer, pp. 17–43.
- (1987). “The principle of contrast: A constraint on language acquisition”. In: *Mechanisms of language acquisition* 1, p. 33.
- Crain, Stephen and Diane Carolyn Lillo-Martin (1999). *An introduction to linguistic theory and language acquisition*. Blackwell Malden, MA.
- Crawford, Jean Lenore (2012). “Developmental perspectives on the acquisition of the passive”. PhD thesis. University of Connecticut.
- Csibra, Gergely, Szilvia Bíró, Orsolya Koós, and György Gergely (2003). “One-year-old infants use teleological representations of actions productively”. In: *Cognitive Science* 27.1, pp. 111–133.

- De Villiers, Jill (1985). "Learning how to use verbs: Lexical coding and the influence of the input". In: *Journal of Child Language* 12.3, pp. 587–595.
- Demuth, Katherine (1989). "Maturation and the acquisition of the Sesotho passive". In: *Language*, pp. 56–80.
- Demuth, Katherine, Francina Moloi, and Malillo Machobane (2010). "3-Year-olds' comprehension, production, and generalization of Sesotho passives". In: *Cognition* 115.2, pp. 238–251.
- Driva, Evangelia and Arhonto Terzi (2007). "Children's passives and the theory of grammar". In: *Language acquisition and development: Proceedings of GALA*.
- Fisher, Cynthia (2002). "The role of abstract syntactic knowledge in language acquisition". In: *Cognition* 82.3, pp. 259–278.
- Fisher, Cynthia, Yael Gertner, Rose M Scott, and Sylvia Yuan (2010). "Syntactic bootstrapping". In: *Wiley Interdisciplinary Reviews: Cognitive Science* 1.2, pp. 143–149.
- Fisher, Cynthia, Henry Gleitman, and Lila Gleitman (1991). "On the semantic content of subcategorization frames". In: *Cognitive psychology* 23.3, pp. 331–392.
- Fodor, Janet D (1985). "Why learn lexical rules". In: *10th Annual Boston University Conference on Language Development, October*. Vol. 25, p. 27.
- Fodor, Jerry A (1970). "Three reasons for not deriving "kill" from "cause to die"". In: *Linguistic Inquiry* 1.4, pp. 429–438.
- Fox, Danny and Yosef Grodzinsky (1998). "Children's passive: A view from the by-phrase". In: *Linguistic Inquiry* 29.2, pp. 311–332.
- Frank, Robert (1998). "Structural complexity and the time course of grammatical development". In: *Cognition* 66.3, pp. 249–301.
- Gleitman, Lila (1990). "The structural sources of verb meanings". In: *Language acquisition* 1.1, pp. 3–55.

- Gleitman, Lila R and Henry Gleitman (1992). "A picture is worth a thousand words, but that's the problem: The role of syntax in vocabulary acquisition". In: *Current Directions in Psychological Science* 1.1, pp. 31–35.
- Gleitman, Lila R and Barbara Landau (1994). *The acquisition of the lexicon*. MIT Press.
- Gleitman, Lila and Elissa L Newport (1995). "The invention of language by children: Environmental and biological influences on the acquisition of language". In: *An invitation to cognitive science* 1, pp. 1–24.
- Goldberg, Adele E (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Goldberg, Adele E, Devin M Casenhiser, and Nitya Sethuraman (2004). "Learning argument structure generalizations". In: *Cognitive linguistics* 15.3, pp. 289–316.
- Golinkoff, Roberta Michnick, Margaret Shuff-Bailey, Raquel Olguin, and Wenjun Ruan (1995). "Young children extend novel words at the basic level: Evidence for the principle of categorical scope." In: *Developmental Psychology* 31.3, pp. 494–507.
- Gómez, Rebecca L and LouAnn Gerken (2000). "Infant artificial language learning and language acquisition". In: *Trends in cognitive sciences* 4.5, pp. 178–186.
- Goodluck, Helen (1981). "Children's grammar of complement-subject interpretation". In: *Language acquisition and linguistic theory* 5, pp. 139–166.
- Goodluck, Helen, Arhonto Terzi, and Gema Chocano Díaz (2001). "The acquisition of control crosslinguistically: structural and lexical factors in learning to licence PRO". In: *Journal of Child Language* 28.01, pp. 153–172.
- Gordon, Peter and Jill Chafetz (1990). "Verb-based versus class-based accounts of actionality effects in children's comprehension of passives". In: *Cognition* 36.3, pp. 227–254.
- Grimshaw, Jane (1990). *Argument structure*. MIT Press.
- Gropen, Jess, Steven Pinker, Michelle Hollander, and Richard Goldberg (1991). "Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure". In: *Cognition* 41.1, pp. 153–195.

- Hart, Betty and Todd R Risley (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.
- Hirsch, Christopher (2011). "The acquisition of raising". PhD thesis. Massachusetts Institute of Technology.
- Hirsch, Christopher, Robin Orfitelli, and Kenneth Wexler (2007). "The acquisition of raising reconsidered". In: *Language acquisition and development: proceedings of GALA*, pp. 253–262.
- Hirsch, Christopher and Kenneth Wexler (2004). "Who seems to rescue raising". In: *Proceeding of generative approaches to language acquisition-North American*.
- (2006). "Children's passives and their resulting interpretation". In: *The proceedings of the inaugural conference on generative approaches to language acquisition-North America, University of Connecticut Occasional Papers in Linguistics*. Vol. 4, pp. 125–136.
- (2007). "The late development of raising: What children seem to think about *seem*". In: *New horizons in the analysis of control and raising*. Springer, pp. 35–70.
- Hollich, George J et al. (2000). "Breaking the language barrier: An emergentist coalition model for the origins of word learning". In: *Monographs of the society for research in child development*, pp. i–135.
- Hsu, Jennifer Ryan, Helen Smith Cairns, and Robert W Fiengo (1985). "The development of grammars underlying children's interpretation of complex sentences". In: *Cognition* 20.1, pp. 25–48.
- Hudson-Kam, Carla and Elissa L Newport (2005). "Regularizing unpredictable variation: The roles of adult and child learners in language formation and change". In: *Language learning and development* 1.2, pp. 151–195.
- (2009). "Getting it right by getting it wrong: When learners change languages". In: *Cognitive psychology* 59.1, pp. 30–66.
- Kline, Melissa and Katherine Demuth (2010). "Factors facilitating implicit learning: the case of the Sesotho passive". In: *Language Acquisition* 17.4, pp. 220–234.

- Landau, Barbara and Lila Gleitman (1985). *Language and Experience*. Harvard University Press.
- Landau, Idan (2013). *Control in generative grammar: A research companion*. Cambridge University Press.
- Legate, Julie Anne and Charles Yang (2007). “Morphosyntactic learning and the development of tense”. In: *Language Acquisition* 14.3, pp. 315–344.
- Lempert, Henrietta (1990). “Acquisition of passives: the role of patient animacy, salience, and lexical accessibility”. In: *Journal of Child Language* 17.3, pp. 677–696.
- Levin, Beth (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago press.
- Levin, Beth and Malka Rappaport Hovav (1992). “The perspective from unaccusativity”. In: *Thematic structure: Its role in grammar* 16, pp. 247–269.
- Lidz, Jeffrey, Henry Gleitman, and Lila Gleitman (2003). “Understanding how input matters: Verb learning and the footprint of universal grammar”. In: *Cognition* 87.3, pp. 151–178.
- Limber, John (1973). “The genesis of complex sentences”. In: *Cognitive development and the acquisition of language*, pp. 169–185.
- MacWhinney, Brian (2000). *The CHILDES project: The database*. Vol. 2. Psychology Press.
- (2004). “A multiple process solution to the logical problem of language acquisition”. In: *Journal of child language* 31.4, pp. 883–914.
- Maguire, Mandy J, Kathy Hirsh-Pasek, and Roberta Michnick Golinkoff (2006). “A Unified Theory of Word Learning: Putting Verb Acquisition in Context”. In: *Action meets word: How children learn verbs*, pp. 364–391.
- Maratsos, Michael (1974). “Preschool children’s use of definite and indefinite articles”. In: *Child development* 45, pp. 446–455.
- Maratsos, Michael, Dana EC Fox, Judith Becker, and Mary Anne Chalkley (1985). “Semantic restrictions on children’s passives”. In: *Cognition* 19.2, pp. 167–191.

- Maratsos, Michael, Roxane Gudeman, Poldi Gerard-Ngo, and Ganie DeHart (1987). "A study in novel word learning: The productivity of the causative." In: *Mechanisms of Language Acquisition*.
- Marcus, Gary F (1993). "Negative evidence in language acquisition". In: *Cognition* 46.1, pp. 53–85.
- Mateu, Victoria Eugenia (2016). "Intervention effects in the acquisition of raising and control: Evidence from English and Spanish". PhD thesis. UCLA.
- McDaniel, Dana, Helen Smith Cairns, and Jennifer Ryan Hsu (1990/91). "Control principles in the grammars of young children". In: *Language Acquisition* 1.4, pp. 297–335.
- McNally, Louise (1992). "VP coordination and the VP-internal subject hypothesis". In: *Linguistic Inquiry*, pp. 336–341.
- Naigles, Letitia G and Edward T Kako (1993). "First contact in verb acquisition: Defining a role for syntax". In: *Child development* 64.6, pp. 1665–1687.
- Naigles, Letitia R (1990). "Children use syntax to learn verb meanings". In: *Journal of child language* 17.2, pp. 357–374.
- (1996). "The use of multiple frames in verb learning via syntactic bootstrapping". In: *Cognition* 58.2, pp. 221–251.
- Naigles, Letitia R and Erika Hoff-Ginsberg (1995). "Input to verb learning: Evidence for the plausibility of syntactic bootstrapping." In: *Developmental Psychology* 31.5, p. 827.
- Nguyen, Emma and Lisa Pearl (2018). "Do You Really Mean It? Linking Lexical Semantic Profiles and the Age of Acquisition for the English Passive". In: *35th West Coast Conference on Formal Linguistics*. Cascadilla Proceedings Project, pp. 288–295.
- Orfitelli, Robyn (2012). "Argument intervention in the acquisition of A-movement". PhD thesis. UCLA.
- Perfors, Amy, Joshua B Tenenbaum, and Elizabeth Wonnacott (2010). "Variability, negative evidence, and the acquisition of verb argument constructions". In: *Journal of child language* 37.3, pp. 607–642.

- Perlmutter, David and Paul Postal (1984). “The 1-advancement exclusiveness law”. In: *Studies in relational grammar 2*, pp. 81–125.
- Pierce, Amy E (1992). “The acquisition of passives in Spanish and the question of A-chain maturation”. In: *Language Acquisition 2.1*, pp. 55–81.
- Pinker, Steven (1984). *Language learnability and language learning*. Harvard University Press.
- (1987). “The bootstrapping problem in language acquisition”. In: *Mechanisms of language acquisition*, pp. 399–441.
- (1989). *Learnability and cognition: The acquisition of argument structure*. The MIT Press.
- Pinker, Steven, David S Lebeaux, and Loren Ann Frost (1987). “Productivity and constraints in the acquisition of the passive”. In: *Cognition 26.3*, pp. 195–267.
- Postal, Paul (1974). “On certain ambiguities”. In: *Linguistic Inquiry 5.3*, pp. 367–424.
- Pye, Clifton, Diane Frome Loeb, and Yin-Yin Pao (1996). “The acquisition of breaking and cutting”. In: *The proceedings of the twenty-seventh annual child language research forum*. Center for the Study of Language and Information Stanford, pp. 227–236.
- Pye, Clifton and Pedro Quixtan Poz (1988). “Precocious Passives (and Antipassives) in Quiche Mayan.” In: *Papers and Reports on Child Language Development 27*.
- Reber, Arthur S (1976). “Implicit learning of synthetic languages: The role of instructional set.” In: *Journal of Experimental Psychology: Human Learning and Memory 2.1*, pp. 88–94.
- Schuler, Kathryn D, Jaclyn E Horowitz, Charles Yang, and Elissa L Newport (2017). “The Sufficiency Principle: Predicting when children will regularize inconsistent language variation.” In: *CogSci*.
- Schuler, Kathryn D, Charles Yang, and Elissa L Newport (2016). “Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient to do so.” In: *CogSci*.

- Sherman, Janet Cohen and Barbara Lust (1993). "Children are in control". In: *Cognition* 46.1, pp. 1–51.
- Shibatani, Masayoshi (1973). "Lexical versus periphrastic causatives in Korean". In: *Journal of Linguistics* 9, pp. 209–383.
- Stevens, Jon Scott, Lila R Gleitman, John C Trueswell, and Charles Yang (2017). "The pursuit of word meanings". In: *Cognitive science* 41, pp. 638–676.
- Talmy, Leonard (1975). "Semantics and syntax of motion". In: *Syntax and semantics* 4, pp. 181–238.
- Tenny, Carol L (1994). *Aspectual roles and the syntax-semantics interface*. Vol. 52. Springer.
- Tenny, Carol Lee (1987). "Grammaticalizing aspect and affectedness". PhD thesis. Massachusetts Institute of Technology.
- Terzi, Arhonto and Kenneth Wexler (2002). "A-chains and S-homophones in children's grammar: Evidence from Greek passives". In: *Proceedings of the North East Linguistics Society*. Vol. 32. 2, pp. 519–538.
- Theakston, Anna L (2004). "The role of entrenchment in children's and adults' performance on grammaticality judgment tasks". In: *Cognitive Development* 19.1, pp. 15–34.
- Theakston, Anna L, Elena VM Lieven, Julian M Pine, and Caroline F Rowland (2001). "The role of performance limitations in the acquisition of verb-argument structure: An alternative account". In: *Journal of child language* 28.1, pp. 127–152.
- Tomasello, Michael (2000). "Do young children have adult syntactic competence?" In: *Cognition* 74.3, pp. 209–253.
- Tomasello, Michael and Michelle E Barton (1994). "Learning words in nonostensive contexts." In: *Developmental psychology* 30.5, p. 639.
- Van Hout, Angeliek (1993). *On unaccusativity: The relation between argument structure and aspect*. Arbeitsgruppe Strukturelle Grammatik, MPG, Berlin.
- (1996). *Event Semantics of Verb Frame Alternations. A Case Study of Dutch and its Acquisition*. TILDIL Dissertations Series.

- Van Hout, Angeliek (2004). “Unaccusativity as telicity checking”. In: *The unaccusativity puzzle: Explorations of the syntax-lexicon interface* 83.
- (2013). *Event semantics of verb frame alternations: A case study of Dutch and its acquisition*. Routledge.
- Vernice, Mirta and Antonella Sorace (2018). “Animacy effects on the processing of intransitive verbs: an eye-tracking study”. In: *Language, Cognition and Neuroscience* 33.7, pp. 850–866.
- Wagner, Laura and Susan Carey (2005). “12-month-old infants represent probable endings of motion events”. In: *Infancy* 7.1, pp. 73–83.
- Wang, Qi, Diane Lillo-Martin, Catherine T Best, and Andrea Levitt (1992). “Null subject versus null object: Some evidence from the acquisition of Chinese and English”. In: *Language acquisition* 2.3, pp. 221–254.
- Wexler, Kenneth and Peter Culicover (1980). *Formal principles of language acquisition*. MIT press.
- Wonnacott, Elizabeth, Elissa L Newport, and Michael K Tanenhaus (2008). “Acquiring and processing verb argument structure: Distributional learning in a miniature language”. In: *Cognitive psychology* 56.3, pp. 165–209.
- Woodward, Amanda L (1998). “Infants selectively encode the goal object of an actor’s reach”. In: *Cognition* 69.1, pp. 1–34.
- Woodward, Amanda L, Jessica A Sommerville, and Jose J Guajardo (2001). “How infants make sense of intentional action”. In: *Intentions and intentionality: Foundations of social cognition*, pp. 149–169.
- Yang, Charles (2002). *Knowledge and learning in natural language*. Oxford University Press on Demand.
- (2005). “On productivity”. In: *Linguistic variation yearbook* 5.1, pp. 265–302.
- (2015). “Negative knowledge from positive evidence”. In: *Language* 91.4, pp. 938–953.
- (2016). *The price of linguistic productivity*. MIT Press.

Yang, Charles and Silvina Montrul (2017). "Learning datives: The tolerance principle in monolingual and bilingual acquisition". In: *Second Language Research* 33.1, pp. 119–144.