NORTHWESTERN UNIVERSITY

Input in the Acquisition of Genericity

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Linguistics

By

Elisa Sneed German

EVANSTON, ILLINOIS

June 2008

ABSTRACT

Input in the Acquisition of Genericity

Elisa Sneed German

This dissertation investigates how a child can acquire the grammar of genericity and what it is that a child actually needs to learn in order to accomplish this task. The first step is to reframe the question asked by the previous literature on the acquisition of generics through the lens provided by linguistic theory. Here, I propose a learning algorithm that presents one way that a child could acquire knowledge of the distinction between generic and referring language. Specifically, I show how a syntactic analysis of genericity, Diesing's Mapping Hypothesis (Diesing 1992), can be combined with a theory of learning that posits a rich system of linguistic representations provided by Universal Grammar (UG) to allow the learner to acquire generic language by using what is in her input.

The findings from two corpus studies show that part of the information the learner needs in order to acquire the grammar of genericity is available in the input. However, the full grammar is not. Instead, the child must rely on innate knowledge of syntactic structure and confirm her hypotheses with the input, rather than learning from it.

A pair of Truth Value Judgment Task (TVJT) experiments provides evidence that children's interpretations go beyond what is in their input and that syntactic structure is used in the interpretation of generic and referring utterances. Further, the second TVJT study provides further experimental support for the VP-internal subject hypothesis and for Diesing's Mapping Hypothesis. Crucially this is demonstrated by priming interpretations of target sentences that are similar to the primes only because they exhibit similar syntactic representations.

Taken all together, the ultimate conclusion from the studies presented in this dissertation shows that UG provides children with a rich grammar that they can use to filter the input in meaningful ways. It is thus that they can demonstrate linguistic sophistication beyond what is in their input. They can apply this to the acquisition of generic and referring language, which serve them well as they learn about the world.

ACKNOWLEDGEMENT

The most difficult part of the dissertation comes now, at the end of it all, when I need to thank everyone who helped me in this endeavor. I have waited as long as possible to write the acknowledgement, not because I haven't had any help, but on the contrary because I have received lots of help and don't possess the words necessary to express my gratitude.

First, I would like to thank my committee members: Jeff Lidz, Chris Kennedy, Sandy Waxman and Stefan Kaufmann. Jeff, my dissertation advisor and chair, has provided tireless support and guidance, even from a distance. He has been generous with his time and advice, as well as with his research participants and research assistants at the University of Maryland. He should win a prize now that his last student at NU is finished. Chris was my first advisor at NU on my first qualifying paper. Although he has since left the department, he has always made himself available to meet and has always provided very insightful comments that have informed every step of the dissertation process. Sandy Waxman has always provided an important perspective from outside of linguistics. Stefan Kaufmann has very patiently discussed issues in modeling and formal semantics when I had questions. His comments, especially on final versions, have this dissertation much better. All of my committee members are all excellent researchers, teachers and mentors and I hope that I have learned some of their skills.

This project started in Jeff and Sandy's seminar on language acquisition and the Waxman-Lidz/Lidz-Waxman lab group provided lots of useful feedback as my project took shape. I thank the many people who were members of that lab group over the years.

I would also like to thank Gregory Ward, Bob Gundlach and Ann Bradlow, chairs of the department in my years at Northwestern. Without all of their hard work with the administration,

a small department like ours would have been a very different place. I especially thank them for their flexibility and willingness to provide opportunities for me to teach while finishing up.

Thanks also go to the RAs in the Project on Child Development, Jane Solomon and Cara Brown (who were excellent puppets) and Micki Sommerman and the teachers and students at Northbrook Community Nursery.

On a more personal note, I would like to thank many of the other students with whom I co-occurred in the department. You are more than colleagues to me and have been good friends to me throughout our time at NU and beyond. In particular, thanks go to Lewis Gebhardt, Carolyn Gottfurcht-Zafra, Ken Konopka and of course to my cohort: Ann Bunger, Deb Gordon, Nick Kaspar, Ralph Rose, Jody Sostrin and John Evar Strid. Some of you contributed more directly to the completion of my dissertation and deserve special thanks. This is especially true of Ann Bunger, who forced me to eat pie the right way and kept me from jumping off the airplane without a pink jumpsuit; Jessica Peterson, who has been a wonderful cosmic advisor; Erin Leddon, who gave me sage advice about telling stories to small children; and Kristen Syrett and Josh Viau, my co-defenders, whose support, advice and breakfast treats in the weeks before our defense day made it bearable.

Last, but certainly not least, my most sincere gratitude goes to my husband James. Without your love, support, encouragement and optimism I would have thrown in the towel in mid-May. And despite advice we were both given not to get involved with another linguist, it's only been a good thing for me.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

Chapter 1

1.0 Introduction

Imagine a child and her caregiver looking through a picture book. They turn to an illustration of two monkeys sitting in a jungle tree eating bananas. The child says, "Monkeys." And the mother says to the child:

1) Monkeys eat bananas.

The next day the child and her father are looking through the same book. They come to the page with the monkeys and again the child says, "Monkeys." The father says to the child:

2) Monkeys are eating bananas.

Both the mother and the father have spoken felicitously, but they have made importantly different assertions. The mother has made an assertion about a generic property of monkeys whereas the father has made an assertion about the two particular monkeys depicted in the book. The problem presented to the learner by examples like (1) and (2) in trying to distinguish generic from referring language would seem to be, in part, the fact that the same string can be used to refer to different things depending on the type of sentence it is used in. For example the string *dog* always refers to something about dogs, but what, exactly that is varies: (3a) refers to a

particular barking dog, (3b) refers either to the particular dog (e.g., *a certain dog*) that is barking, or to a non-specified dog and (3c) refers to a generic property of the kind [dog].

3) a. The dog is barking.

b. A dog is barking.

c. Dogs bark.

While some element of the meaning of *dog* may be constant across these examples, what's being referred to is determined by the whole NP (e.g., *the dog*, *a dog*, *dogs*) at a minimum or possibly by the whole sentence. The problem of determining an NP's intended interpretation is compounded by the observation that a given NP form is often compatible with both generic and existential or referential interpretations depending on the sentence it occurs in, as shown in (4) and (5). Moreover, there is no unique NP form that indicates genericity.

4) a. The cow has four stomachs.

b. A camel spits three times before sleeping.

c. Tasmanian tigers are extinct.

5) a. The cow is eating.

b. A camel spit three times.

c. There are Tasmanian tigers in the natural history museum.

Previous research has argued that the many-to-many mapping between NP form and NP function with respect to generic and referring language presents a significant learning problem. This dissertation will show that, in fact, it does not; nonetheless, sentences like those in (1) – (5) raise a number of questions about what a learner understands when presented with utterances like these. Does she notice the morphosyntactic differences? How does a child make use of the information in the sentence to guide his interpretation of these utterances? Most importantly, what is it that a child actually needs to learn to understand the difference between utterances like those in examples (1) and (2) or (4) and (5)?

This last question is the focus of this dissertation. In order to answer the question, this dissertation unifies disparate lines of research from different disciplines that do not always use the same terminology or operate on the same background assumptions. I will argue that children can solve the problem of learning about generic and referring language by taking advantage of an innate knowledge of the syntax-semantics map for genericity. Assuming a rich innate structure for the learner in no way trivializes learning. On the contrary, it makes learning possible by permitting the learner to acquire grammatical knowledge on the basis of the input. It allows the learner to filter the input in meaningful, informative ways. By restricting the domain to a subset of the widely varied phenomena that have been called generics, I provide a model of how, given the input, a child can acquire the grammar of genericity. In particular, I will show how one linguistic analysis of genericity (Diesing's (1992) Mapping Hypothesis) can be combined with a theory of learning (which exploits the use of the input within the rich system of linguistic representations provided to the learner by Universal Grammar) in order to show how children acquire the right grammar.

In the remainder of this chapter, I will begin by discussing the phenomena that have been described as generic and the problems that are met when trying to pin down what it means to be "generic". Then I will discuss the grammar of genericity. This will include a general discussion of quantification, as well as a discussion of the interpretation of indefinites and finally the use of the Mapping Hypothesis (Diesing 1992) as a basis for a learning algorithm. I will then discuss what children would have to know in order to make use of the learning algorithm and how the previous literature on the acquisition of generics bears on this question. Finally, as a preface to the experimental chapters that address open issues in the literature, I will show how treating genericity as a property of sentence grammar, rather than as a property of NPs, constrains the hypothesis space to one that children could use to learn to distinguish generic and referring language. Briefly, children know which syntactic representations map onto which semantic representations by virtue of the mapping hypothesis being part of their grammars. They do not need to learn this from the input. Children need the input, but to determine which morphological forms of NPs are indefinites, definites, or quantificational NPs.

Note to the reader

The reader should be alerted that this is not a dissertation on generics. It is a dissertation on the acquisition of generics that will, by examining only a subset of the phenomena that have been called generics, answer the question of how children can gain insight into the difference between generic and referring language. Therefore, I will not discuss generics in the detail that should be expected of a dissertation on that topic. I will also remain agnostic about many issues that are part of genericity. I have chosen a model of generics that allows me to develop a theory of the acquisition of generics but I will not critique the model or discuss any of its imperfections as a

theory of generics because that is beyond the scope of this dissertation. My goal in writing this dissertation is to explain one way that children could acquire knowledge of a particular area of the grammar that is not, on the surface, transparent to the language learner. To do this, I will connect the findings in the literature on the acquisition of generics to the theoretical literature on generics. I will argue that by bringing the acquisition literature in line with the theoretical literature, what seemed like a difficult or impossible learning problem is in fact manageable. I will discuss the linguistic input children get that can help them on their way to an adult grammar, what their grammars must already possess for them to make use of the input and how children are able to go beyond the limited input they receive to progress toward the fully fledged adult system in their language.

My focus will be on English and unless otherwise specified, I will be talking about English alone, although this dissertation will also touch on cross-linguistic issues in this domain.

1.1 Generics: A preliminary description

Generic utterances are statements that allow speakers to generalize away from particular individuals, events or facts. Rather, they allow us to make statements about the shared properties of such a group (e.g., of individuals (6a), particular types of events (6b)[1], or facts (6c)).

6) a. Dogs bark.

b. John smokes a cigar after dinner.

c. A potato contains vitamin C, amino acids, protein and thiamin.

[1] Crucially, the type of event in question here involves a generalization about after-dinner events that contain John and a cigar. It does not convey a generic property of John or of cigars.

Generic utterances involve properties that are not only descriptive of the category, but are "law-like", "essential" or "normative" to the category (Dahl 1975[2]). They are used to make statements about the category (of individuals, events or facts) as a whole, like the utterance in (1), not to describe individual members of the category or a particular set of category members (Lyons 1977), like the utterance in (2). Generics are also unique in the way they are used to make generalizations because they seem to both refer to the whole category (unlike *some*) and yet, necessarily not the whole category because they allow for exceptions (unlike universally quantified statements with *all*). In fact, generic statements are generally considered to be true even if they are not true for most of the exemplars within the category (e.g., *Birds lay eggs* in (7a) is true only of a subset of birds).

7) a. Birds lay eggs.

b. Some birds lay eggs (but some do not).

c. All birds lay eggs.

As Chierchia (1998) suggests with his description of kind-NPs[3], what counts as a "generic" "… is not set by grammar, but by the shared knowledge of a community of speakers. It thus varies, to a certain degree, with the context, and remains somewhat vague" (p. 348). This description of generics is particularly insightful because it captures the notion that there is

[2] Dahl also uses the term "nomic" to refer to generic utterances, but is careful to distinguish this from "gnomic". In fact, he explicitly avoids discussion of how gnomic utterances and generic utterances are similar (1975: p.100); however, others (e.g., Lyons 1977) discuss the significant overlap between the timelessness and omnitemporality of gnomic utterances and generic utterances.

[3] Briefly, a kind-NP is an NP that refers to a kind as a whole. By Chierchia's definition, what counts as a "kind" is defined by the community of speakers. Unlike a great deal of the literature on generics and the acquisition of generics, this characterization of kind-NPs seems to be applicable to the entire range of linguistic phenomena that have been called generics.

something conceptually uniform about generic utterances (namely, that they are statements that allow speakers to generalize away from particular individuals, events or facts). Crucially, though, generics are not the uniform linguistic phenomenon that much of the previous literature on their acquisition has assumed.

1.2 Problems defining genericity

One potential source of confusion over how to look at the acquisition of genericity may derive from the fact that what it means to be generic is somewhat varied. In some ways, genericity seems to be a property of the NP or DP, but in others it seems to be a property of the whole sentence. This is because the term genericity can be used to refer to (at least) two very different phenomena (Krifka et al. 1995). For the purposes of this dissertation, I will assume that one type of genericity is reference to a kind; the other is reference to a regular property that characterizes individuals, events or facts. The motivation for grouping both types of genericity together is clear: both pertain to generalizations, not to individual exemplars (whether these exemplars are particular objects or particular situations). However, these two phenomena are distinct and should not be subsumed under a single label as they have been in the literature on the acquisition of generics[4].

[4] The semantics literature has long recognized the two-way distinction I mention here, as well as many more subtle differences within these two categories. Carlson (1977b) treats both these types of genericity (although without explicitly labeling the distinction "generic NPs" and "characterizing sentences"), but other previous research has focused on one or the other. The interested reader is directed to Bacon (1973a, b) for a discussion of generic NPs, Kleiber (1989a) for a discussion of generic "*le*" in French, and to Ojeda (1991) for a discussion of the ambiguity of the English definite singular between a definite description and a generic. Burton-Roberts (1976) focuses on the indefinite singular "generic" and distinguishes it from other generic NP forms, but in so doing ultimately focuses on characterizing sentences. Likewise, Dahl (1975) focuses on what he terms "generic tense" and briefly relates them to "indefinite generic noun phrases" like *beavers* and *a beaver*. A more exhaustive list of resources can be found in Krifka et al. (1995).

The first type of genericity mentioned in this section is illustrated in the sentences below ((8a) – (8c) from Krifka et al.) and (8d), where the underlined NPs refer to the kind [potato] or [rice]:

8) a. The potato was first cultivated in South America.

b. Potatoes were introduced into Ireland by the end of the 17th century.

c. The Irish economy became dependent on the potato.

d. Rice was first cultivated in India.

In the four examples in (8), the underlined NPs do not denote any specific, individual potatoes, or make generalizations about the individual potatoes (or rice) that make up the class [potato] (or [rice]); rather, these sentences are about the kind [potato] as a whole. These are KIND-NPs and as such, they are also generic-NPs.

Kind-NPs are distinguished in the literature from OBJECT-NPs which denote objects that are of the kind[5]. Many predicates can take either kind- or object-NPs as their arguments, but there is a certain class of predicates, KIND-PREDICATES, that require reference to the kind as a whole (e.g., *die out*, *be extinct*, *be widespread*). For instance, predicates like *be extinct* and *die out* can only apply to the entire kind because while a particular object can be a member of a species that is extinct and can itself die, the object itself cannot be extinct or die out. Consider the following examples in (9) and (10), where an asterisk next to an utterance indicates its ungrammaticality on the particular reading under discussion.

[5] Throughout this dissertation, I will use the term OBJECT-NP uniquely in this sense. Direct objects of verbs will be referred to as such and their syntactic position will be referred to as direct object position.

9) a. The Tasmanian tiger is extinct.

b. Tasmanian tigers are extinct.

c. * A Tasmanian tiger is extinct.

10) a. The Sonoran pronghorn is dying out.

b. Sonoran pronghorns are dying out.

c. * A Sonoran pronghorn is dying out.

d. A pronghorn, namely the Sonoran pronghorn, is dying out.

e. A Sonoran pronghorn is an ungulate, but not a true antelope.

The ungrammaticality of the examples in (9c) and (10c) provides evidence that in English, kind-NPs can only take the morphological form of definite singular, bare plural or, as evidenced by (8d), mass nouns. Notably absent from the semantic category of kind-NP is the indefinite singular NP, which has often been treated like a generic-NP (e.g., Bacon 1973b, Dahl 1975, Burton-Roberts 1976, Lyons 1977, Gelman & Tardif 1998). However, an indefinite singular NP cannot be treated as a kind-referring expression unless it receives a taxonomic reading, as in (10d)[6]. When the genericity of a sentence is a feature of the whole sentence and not limited to an NP in it, the indefinite singular can still be used to express regularities as in (10e). The sentence in (10e) provides a preview of the second type of genericity, discussed below.

The second type of genericity involves a proposition that expresses a general property of individuals, episodes or facts. As in the above examples with NPs, the type of proposition under

[6] Examples (9c) and (10c) would also be acceptable on a taxonomic interpretation if the speaker were referring to a particular subspecies of the kind [Tasmanian tiger] or to a subspecies of (the subspecies) [Sonoran pronghorn].

i. A Tasmanian tiger, namely Dickson's Thylacine, is extinct.

ii. A Sonoran pronghorn, namely the "Nogales" pronghorn, is dying out.

discussion does not refer to any specific instance; rather, it serves as a generalization about the individuals, episodes or facts. The sentences in (11) exemplify genericity through the whole utterance, not, as we saw in the examples in (8), through the subject NP in them. Krifka et al. (1995) call sentences like these CHARACTERIZING SENTENCES because they allow any type of NP to be interpreted generically and can even be used to make characterizing statements about individuals using a proper name. The fact that the most likely interpretation of the subject NP seems to vary from sentence to sentence is further support for the claim that the sentence itself is what gives rise to genericity in these cases. Note the variation in what the subject NP denotes in the sentences in (11): in (11a), it denotes a particular individual; in (11b) and (11c), it seems to pick out a typical exemplar of a psychologically defined kind or a group of typical exemplars of a psychologically defined kind, respectively; in (11d) and (11e), it seems to denote to the kind as a whole.

11) a. Ann bakes cookies on the weekend.

b. A cookie contains butter, sugar, flour, baking soda and vanilla extract.

c. Cookies contain butter, sugar, flour, baking soda and vanilla extract.

d. Milk is healthy.

e. The alpaca has a cheery disposition.

Characterizing sentences stand in contrast to PARTICULARIZING SENTENCES, which express a proposition about a particular event. Particularizing sentences can be interpreted generically only when they contain a kind-referring-NP. This is true of the sentences in (8), in

which each utterance can be said to be about a particular occurrence, but makes a generalization about the kind [potato] as a whole[7, 8].

It should now be clear to the reader that while, intuitively, there is a conceptual category of "generic statements", there are different reasons for an utterance to be classified as belonging to that category: an utterance may be so categorized either because it contains a kind-NP *or* because it is a characterizing sentence, independent of any of the NPs that are in it. The previous literature on the acquisition of genericity (e.g., Pappas & Gelman 1998, Gelman & Tardif 1998, Gelman et al. 1998, Hollander et al. 2002, Gelman et al. 2002, Gelman & Raman 2003) has not made this distinction, confusing the more uniform conceptual category of "generic" for a uniform linguistic category. The importance of distinguishing these types of genericity may be further confounded by the fact that kind-NPs can occur in characterizing statements as in (11c), (11d) and (11e) and in the examples in (12) below. In such instances, there are two sources of genericity. Consider the following statements where characterizing sentences and kind-referring NPs are combined in one sentence. In the examples in (12), the characterizing sentence serves to generalize over facts about the kind itself (in (12a) and (12b)) and in (12c) the characterizing sentence generalizes over individual eggplants as they are typical of the kind.

12) a. The eggplant is higher in nicotine than any other edible plant.

b. Eggplants are higher in nicotine than any other edible plant.

c. An eggplant is higher in nicotine than any other edible plant.

---

[7] Note that thus far, I have only discussed the generic interpretation of subjects. Generic interpretations of direct objects are somewhat more complicated and will be discussed in greater detail later.

[8] Krifka et al. (1995) provide a series of tests that distinguishes kind- from non kind-NPs and characterizing from particularizing sentences. The reader is directed to *The Generic Book*, Ch. 1 for a detailed discussion of the different tests.

Considering the interaction between kind-NPs and characterizing sentences illuminates the subtleties of the issue. While both the definite singular and the bare plural are considered to be kind-NPs, they allow different interpretations. The definite singular is ambiguous between denoting the kind as an undifferentiated whole and denoting a particular individual, as shown in the sentence in (12a), where *the eggplant* can either denote the kind or it can denote a single, particular eggplant that is sitting on my kitchen counter (i.e., *that* eggplant). On the other hand, the bare plural is ambiguous between denoting the kind as an undifferentiated whole, and denoting a plurality of instances of the kind and making a generalization about them[9]. Crucially, for the bare plural, the truth conditions seem to be identical whether the NP denotes the kind as a whole or typical instances of the kind, while the truth conditions are not the same for the two interpretations available for a definite singular NP. While it may be accepted as true that the kind [eggplant] is higher in nicotine than any other edible plant, the eggplant that is sitting on my counter may be a particularly poor specimen, and not contain any nicotine whatsoever. In this case, (12a) is still true on its generic interpretation, which allows for exceptions, but false on the particular (i.e., *that* eggplant) interpretation. The truth conditions are the same for the two denotations of the bare plural because in uttering (12b), the speaker is making a claim about the kind, or about typical instances of the kind.

Another distinction between the two types of kind-NPs arises with their available interpretations in particularizing sentences. The ambiguity described above for a definite singular

---

[9] This is particularly clear if we change the predicate to one for which plurality matters as in the example below.

i. Eggplants are a member of the nightshade family.
ii. Eggplants are members of the nightshade family.

The predicate *be a member of* is felicitous when *eggplants* denotes the kind (i), whereas the very similar *be members of* is felicitous when the bare plural makes generalizations about instances of the kind (ii). See also the discussion in Krifka et al. regarding the felicity of bare plural and definite singular NPs and the infelicity of indefinite singular NPs in characterizing sentences with collective predicates (1995: p. 89-90).

NP does not arise when it is used in a particularizing sentence. In example (13a) (=(8b)), the definite singular NP must receive a kind-interpretation or the utterance is just as ungrammatical, as illustrated in (13b), as a kind-predicate with an indefinite singular NP (13c and 13d). This is because an individual exemplar cannot be cultivated (whether it is denoted by a definite or an indefinite NP), but a kind can. Conversely, bare plural subjects are again ambiguous between denoting instances of the kind and denoting the kind as an undifferentiated whole. The truth conditions of the two interpretations of the bare plural NP in (13e) are indistinguishable.

13) a. The potato was first cultivated in South America.

b. * The potato (i.e., that exemplar) was first cultivated in South America.

c. * A potato was first cultivated in South America.

d. * A potato is widespread.

e. Potatoes were first cultivated in South America.

In what follows, I will limit my discussion to characterizing sentences like the examples in (14). These examples were produced by English-speaking mothers in Gelman & Tardif (1998)[10], which was a series of studies designed to examine use of generic and nongeneric language by English- and Mandarin-speaking mothers.

14) a. A kite goes flying in the sky.

b. Bunny rabbits don't make noises, do they?

c. Do you like buttons?

[10] This series of transcripts is available as the Tardif corpus in the CHILDES database (MacWhinney 2000).

These are the type of generic utterances that the previous acquisition literature has focused on, though it has done so without calling them "characterizing" utterances or distinguishing them from generic utterances that contain kind-NPs. As my analysis shows, however, it is important to make a distinction between the two types of genericity discussed above because the two represent very different linguistic phenomena. Crucially, unlike kind-NPs, object-NPs are not generic expressions on their own. They are interpreted generically only if they occur in characterizing sentences.

Throughout this dissertation, I follow Krifka et al. (1995) in adopting a very broad definition of characterizing sentences that subsumes the different kinds of generic sentences that have been discussed in the literature. It is important to note that I am simplifying a great deal by not addressing the differences between "descriptive" and "normative" statements (see Dahl 1975) or among "habitual", "existential", "universal" and "occupational" generics (see Lawler 1972). However, these differences are beyond the scope of this dissertation. For the purposes of the discussion here, I will assume that all these types of generic sentences involve quantification over situations, which will be discussed in §1.3.

I will not discuss genericity via kind-NPs in great detail. The use of the definite singular NP as a generic expression has been shown to be very rare, if present at all, in speech to children (Gelman et al. 1998) and in adult-adult speech, it seems to occur only in formal, academic contexts. These factors make it likely that the use of a definite singular NP as a generic expression is learned much later in language development, and perhaps only prescriptively. As noted above, the interpretation of the bare plural as a kind-NP is not separable from its interpretation as an object-NP in a characterizing sentence. Further, the input to children from the previous literature on the acquisition of generics does not include the use of particularizing

sentences (e.g., the sentences in (8a-d)) that receive their generic interpretation because they contain a bare plural kind-NP. This is true for the speech of caregivers (Pappas & Gelman 1998, Gelman & Tardif 1998, Gelman et al. 1998), the speech used by experimenters in experimental tasks in the lab (Hollander et al. 2002, Gelman et al. 2002, Gelman & Raman 2003), and transcripts of child-directed speech (Brown 1973, Gelman & Tardif 1998, MacWinney 2000). Nor do these sources include the use of kind-predicates in child-directed speech[11]. Since there are no kind predicates in the sample of child-directed speech analyzed in this dissertation, we should assume that the bare plural NPs in speech to children are object-referring NPs in characterizing sentences and overlook the analysis of kind-NPs.

1.3 The grammar of genericity

The following proposal is based on work in linguistic theory that has received support from facts of the adult grammars of several languages. The account I propose brings together several lines of research that have not been previously linked. However, when these pieces are examined and connected, they point to a system of acquisition of generics that offers a child a one-to-one mapping from structural position to interpretation. This is precisely the type of system that a child could learn from. As noted above, in the proposal that follows, I will not discuss the acquisition of the definite singular generic. I will discuss, instead, how children could learn the difference between the generic and referring language that they actually do hear. This account appeals to linguistic theory on quantification in general, and more specifically to the theory of the quantification and interpretation of indefinite NPs.

[11] Recall that the only way to unequivocally determine that an NP is kind-referring is through its grammatical occurrence with a kind-predicate (see examples (9) and (10) in §1.2).

1.3.1 Quantification

Quantification allows speakers to express generalizations over subsets of a set. The previous linguistics literature has identified two primary varieties of quantification (summarized in Partee 1995, cited as Partee et al. [1986]), each of which involves quantification over different elements of a sentence. Determiner quantification occurs when the quantifier (e.g., *every*, *some*, *two*) is the head of the DP, as in (15). Adverbial quantification, on the other hand, does not involve determiners directly; rather, in this case the quantificational adverb (e.g., *always*, *usually*, *rarely*) modifies a clausal projection as in (16).

15) Every runner finished a race.

16) Usually runners have callused feet.

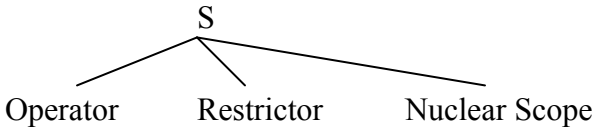
Both these types of quantification are restricted quantification in which the domain of the quantifier is established by a restrictive clause. Every quantificational expression (e.g., *every*, *two*, *always*, *usually*) denotes a relation between two sets (the set denoted by the restrictor and the set denoted by the nuclear scope). In the examples in (15) and (16), the restrictor is the set of runners and the nuclear scopes are the set of individuals that finished a race and have callused feet, respectively. The quantificational expressions tell us that in (15), *all* the individuals that are runners are members of the set that finished the race and in (16) that, roughly, *most* of the individuals that are runners are members of the set that have callused feet. In the case of adverbial quantification, the quantificational adverb (Q-adv) can also restrict an abstract event variable introduced as an invisible argument of episodic (non-stative) predicates (Davidson

1967). A great deal of research has argued in favor of the presence of such a spatiotemporal argument that ranges over locations and times or situations (e.g., Lawler 1972, Spears 1974, Kratzer 1995; Schubert & Pelletier 1989). In the case of (17), this abstract variable would be over times and locations that include Elaine and Michelle participating in a 10k race[12]. I will refer very generally to such variables as situation variables.

17) Usually Elaine beats Michelle in a 10k race.

Heim (1982)'s tripartite structure, represented below, captures what is common to the semantic representations of both determiner and adverbial quantification. Included in Heim's (1982) analysis is an implicit existential (∃) quantifier that adjoins to the nuclear scope and unselectively binds any unbound variables in the nuclear scope to prevent the occurrence of unbound variables in the logical representation. If there is no quantificational operator, there is no restrictor, only a nuclear scope. In such a case, the unbound variables would all be bound by the implicit ∃-quantifier (see example (23) in §1.3.2 below). In this respect, ∃-closure is the default. Both the interpretation of quantificational expressions, through Quantifier Raising (QR), and ∃-closure occur at an abstract level of grammatical representation called Logical Form (LF) (May 1985).

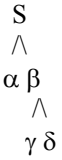
18) Operator [[restrictor] [nuclear scope]]



[12] Example (6b) is another such example.

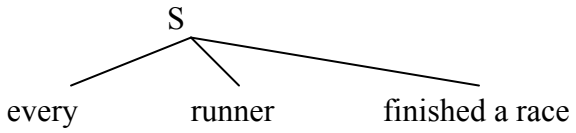
In a traditional GB model there are three levels of syntactic representation: deep structure (D-structure), surface structure (S-structure) and LF. D-structure, roughly, is the level of representation where the thematic relations are specified. D-structure is the input to the transformational component of the grammar, the output of which is S-structure. The S-structure of an utterance contains the syntactic information necessary to interpret the utterance. LF exists as an abstract level of representation, from which the semantic interpretation of an utterance is derived. LF has been postulated to account for the influence of syntactic structure on semantic interpretation. For instance, when quantified NPs (QNP) undergo QR, they raise from their S-structure position and adjoin to IP. Thus the same surface string at S-structure can represent two semantic interpretations, which are distinguished at LF by their scopal relations[13]. The semantic representations below reflect the LF of the utterance. LF is an important part of the analysis of generics and we will return to it later in more detail. Let us now demonstrate how a sentence like (15) would be represented in Heim's (1982) tripartite structure. The logical representation of sentences with implicit ∃-closure will represented by putting the ∃-quantifier inside the brackets representing nuclear scope (as in (15)).

[13] Scopal relations depend on a syntactic relation called "c-command". One element of a sentence can be said to have scope over the other elements in its c-command domain. The c-command domain is defined as follows: α c-commands β iff neither α nor β dominates the other and the lowest branching node that dominates α also dominates β. Graphically, this is represented in (i) where α c-commands β, γ and δ, but only β c-commands α.

i)


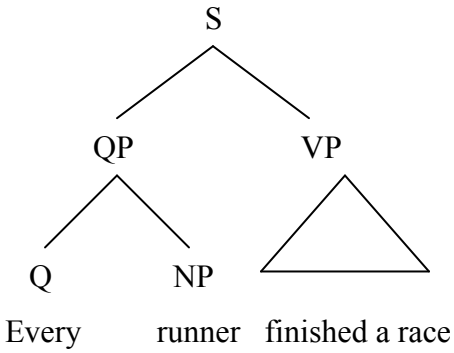
15) Every runner finished a race.

$\forall_x$ [[x is a runner] [$\exists_y$: y is a race & x finished y]]



In the case of determiner quantification, the division between the restrictor and the nuclear scope is actually expressed directly in the syntax. The quantifier is the operator, the first argument (i.e., the restrictor) is the sister of Q° and the second argument (i.e., the scope) is the sister of the whole QP, which is the VP.

19)



The details are somewhat different for adverbial quantification, but the same basic relationships hold. Following the analyses of Lewis (1975) and Heim (1982), the Q-adv (like the quantifier in determiner quantification) is the operator and takes two sentential complements as its arguments. Lewis and Heim represent the internal argument of the Q-adv as an "if-clause" (explicit or implicit) that functions to restrict the domain of quantification and the external argument as the matrix clause. Note the following representation of the logical form of a Q-adv

in (20a), with the particular representation for usually in (20b) adapted from Heim's representation for *in most cases*. The representation in (20a) is a way of presenting the semantic representation in (18) specifically for Q-advs, reflecting the fact that their two arguments are sentential. Like QNPs, Q-advs also take scope over the elements they c-command at LF[14].

20) a. Q-adv (φ,ψ)

b. *usually* (φ,ψ)

"usually (φ,ψ)" is true iff more than half of the ways of assigning values to the free variables that make φ true also make ψ true.

Each of the sentences discussed so far has included only one quantificational element. It is obviously possible that utterances with multiple quantificational NPs occur. In such cases, the ordering of the quantificational operators in the LF structure indicates the semantic interpretation. The semantic representations of the utterance below would be represented hierarchically in the syntax with *every* occurring higher in the tree structure than *some* in (21a) and vice versa in (21b).

---

[14] There are a number of factors, including even intonational features like stress or pitch accent placement that can influence what falls into the restrictor of the quantifier or the nuclear scope. Different semantic interpretations are represented in the syntax by the interpretation of the adverbial quantifier in different locations in the clausal structure at LF. For instance, the topic falls in the restrictor while focused material falls into the nuclear scope. Note the following example from Partee (1991), where boldface indicates pitch-accented material and *s* indicates a situation variable that is quantified over by the operator.

i. Mary always took **John** to the movies.
Always$_s$ [[Mary took *x* to the movies at *s*] [Mary took John to the movies at *s*]

ii. Mary always took John to the **movies**.
Always$_s$ [[Mary took John to *x* at *s*] [Mary took John to the movies at *s*]]

An example like (17) would be sensitive to the same focus effects depending whether "Michelle" or "10k" is accented. It is also possible that rather than binding the abstract variable over situations containing Elaine and Michelle that involve 10k races, the Q-adv could bind the variable introduced by the indefinite "a 10k race" resulting in the following semantic representation:

iii. Usually$_x$ [[(if) *x* is 10k race] [Elaine beats Michelle in *x*]]

21) Every student read some book.

a. $\forall_x$ [[*x* is a student] [$\exists_y$: *y* is a book & *x* read *y*]]

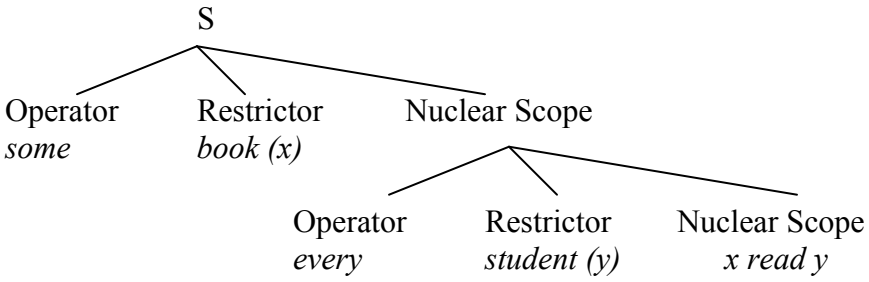
For every student, that student read some book.

b. $\exists_x$ [[*x* is a book] $\forall_y$ [*y* is a student & *y* read *x*]]

There is a particular book that every student read.

The tripartite representation for sentences like (21a) is the same as the representation illustrated in (15). However, the representation of a sentence like (21b), where both of the quantificational elements require a restrictor, requires a slightly different tripartite structure. Consider the case of (21b), where *some* takes scope over *every*. Here, the existential does require a restrictor[15], but we would still want to say that *student* is the restrictor of *every*[16]. In this case the nuclear scope of the first operator consists of another operator, restrictor and nuclear scope as in (22).

22) Operator [[restrictor] [nuclear scope]]



[15] I will not discuss wide-scope indefinites in more detail here, but see Kratzer (1998).
[16] For this reason, in the representation, *every* is not inside the brackets to represent a nuclear scope in (21b) the way *some* is in (21a) or in (15).

Consider again example (17). The truth of this utterance depends only on whether the particular variable assignments that satisfy φ (i.e., *x* is a 10k race and Elaine and Michelle compete in *x*) are variable assignments which also satisfy ψ (i.e., Elaine beats Michelle at *s*).

17) Elaine usually beats Michelle in a 10k race.

usually (φ,ψ)

usually$_s$ [[$\exists_x$: (if) *x* is a 10k race and Elaine and Michelle compete in *x*] [Elaine beats Michelle at *s*]]

This semantic representation is analogous to the semantic representation of determiner quantification in (15).

1.3.2 Semantic interpretation of indefinites

The above discussion of quantification is particularly relevant for the interpretation of a class of nominal elements called INDEFINITES because indefinites receive their quantificational force from other elements in the sentence (Kamp 1981, Heim 1982). It has long been noted in the literature (e.g., Lewis 1975) that simple existential quantification cannot account for the range of interpretations that are available for indefinites, including (in English) indefinite singular NPs (e.g., *a mouse*) and bare plural NPs (e.g., *mice*). For instance, consider the sentence in (23),

where the NP seems to refer to a single mouse (and the direct object NP to a single scrap of bread).

23) A mouse ate a scrap of bread.

Now consider the examples in (24) where the indefinite is interpreted as ranging over mice. The Q-adv in (24a) indicates that most mice have this property whereas the adverb in (24b) indicates that this property is true of relatively few mice[17].

24) a. A mouse usually hates cats.

b. A mouse rarely hates crumbs.

Kamp (1981) and Heim (1982) argued that these data show that indefinites are non-quantificational and introduce variables that have no quantificational force of their own. Instead, indefinites pick up their quantificational force when they are bound by other operators in the sentence, such as Q-adv (as in (24)) or ∃-closure (as in (23)). More concretely, the indefinite introduces a variable which is interpreted as an existential when it has been quantified through ∃-closure as in (23) or which is interpreted as the restrictor of operator. In the examples in (24), this is the Q-adv *usually* or *rarely*. It is now probably very clear how the facts on quantification that were discussed in §1.3.1 are relevant for the interpretation of indefinite NPs.

[17] It is important to note that (a) need not be interpreted as meaning "there is a (particular) mouse who usually hates cats" and (b) need not be interpreted as meaning "there is a (particular) mouse who rarely hates crumbs" simply because they contain the same indefinite singular NP *a mouse*. Note, however, that such readings are available if the ∃-quantifier scopes over the Q-adv.

Every sentence can be mapped into a logical representation of the type introduced in §1.3.1. However, not every sentence will be divided into a restrictive clause *and* a nuclear scope. Recall that if there is no quantificational element in the sentence, then no restrictive clause will be formed and the logical representation will only be a nuclear scope. All the variables introduced by indefinites will be bound by the implicit ∃-quantifier, which unselectively binds all unbound variables in the nuclear scope. This is the case for a sentence like (23).

23) A mouse ate a scrap of bread.

[$\exists_{x,y}$: x is a mouse & y is a scrap of bread & x ate y]

When a sentence does contain an operator (e.g., *every* in (15) or *usually* in (24a)), the operator provides the quantificational force and the restrictor specifies the set that the operator quantifies over. In other words, the restrictor stands in the particular relationship to the nuclear scope that is indicated by the quantifier. A particular sentence will be true only if the sets actually stand in that relation. In this case, the tripartite form in example (18) can be used to represent the semantic configuration of a sentence with an indefinite. Recall the semantic representation of (15), where only the variable introduced by *a race* is bound by existential closure.

15) Every runner finished a race.

$\forall_x$ [[x is a runner] [$\exists_y$: y is a race & x finished y]]

Semantic representations of the sentences in (24) containing indefinites that are bound not by a quantifier but by a Q-adv are provided below.

24) a. A mouse usually hates cats.

usually$_x$ [[*x* is a mouse] [*x* hates cats]]

b. A mouse rarely hates crumbs.

rarely$_x$ [[*x* is a mouse] [*x* hates crumbs]]

Thus far, the operators used in the examples here have been overt, with the exception of the implicit ∃-quantification that occurs via ∃-closure. However, ∃ is not the only covert operator that can quantify over indefinites. The semantics of genericity that I will adopt here is based on the literature on the quantification of indefinites that has been discussed in this section. I will assume, following the research of Lewis (1975), Kamp (1981), Heim (1982) and Diesing (1992) that the generic operator GEN, though covert, behaves like any of the overt quantificational operators. Unlike implicit ∃-closure, but like Q-adv, the generic operator induces the formation of a restrictive clause. Consider the following example.

25) A bird usually has wings.

usually$_x$ [[*x* is a bird] [$\exists_y$: *y* are wings & *x* has *y*]]

The analysis provided here for adverbial quantification by *usually* can be extended very straightforwardly to the abstract generic operator, GEN. Quantification by the generic operator

proceeds much the same way. As in the examples of adverbial quantification above, the truth of (26) depends only on whether the particular variable assignments which satisfy φ (i.e., be a bird) are the variable assignments that also satisfy ψ (i.e., have wings).

26) A bird has wings.

GEN_x [[x is a bird] [$\exists_y$: y are wings & x has y]]

The analysis is the same if the indefinite is a bare plural NP.

27) Birds have wings.

GEN_x [[x is a bird] [$\exists_y$: y are wings & x has y]]

Kamp (1981) and Heim (1982) deal only with what happens to indefinites in the semantics. Diesing (1992), however, builds upon their accounts, proposing the Mapping Hypothesis (MH), which defines the role of the syntax in determining how an indefinite is interpreted – whether it is part of the restrictor of a quantifier or part of its scope. Diesing's (1992) claim is that syntactic structure (word order and hierarchical structure) does play a role in the semantic interpretation of NPs and that syntactic structure can be used to determine which NPs in a sentence fall into which semantic partition (i.e., nuclear scope or restrictive clause) in a Kamp-Heim system of logical representations. The MH is concerned primarily with interpretations of subjects; however, Diesing does treat objects separately and they will be discussed later. Although she does discuss quantification with overt Q-adv, Diesing's focus is on

quantification by the covert generic operator and via ∃-closure, and how quantification by these operators is dependent on syntactic position.

28) a. Birds have wings.

b. Birds are chirping in the tree outside my window.

Likewise, while Diesing briefly discusses quantification of the variable introduced by indefinite singular NPs, her focus is on bare plural NPs that were observed by Carlson (1977b) to be ambiguous between a generic and an existential interpretation as illustrated in (28).

1.3.3 The Mapping Hypothesis

In her analysis, Diesing (1992) assumes a traditional GB model with the three levels of syntactic representation discussed in §1.3.1: D-structure, S-structure and LF. These levels of representation are relevant for Diesing's analysis because she assumes that there are two subject positions in the syntax: [Spec, IP], where the subject appears in the surface representation of an utterance (and where it gets case) and [Spec, VP], the base generated VP-internal subject[18], where the subject's relationship to the verb is specified (through θ-role assignment at D-structure). These two structural positions correspond to the two parts of the semantic representation: the restrictor and the nuclear scope. Diesing also assumes the existence of LF as a level of interpretation, where movement rules follow Chomsky's (1981) Move-α schema.

[18] A great deal of research provides empirical support for the VP-internal subject hypothesis including Kitagawa (1994) for subjects in English and Japanese, Contreras (1987) for small clauses in Spanish and English, Kuroda (1988/1993) for agreement in Japanese and English, Sportiche (1988) for floated quantifiers, Koopman & Sportiche (1991) for position of subjects in several languages, McCloskey (1996) for subject positions in Irish, and Musolino et al. (2000) for the interpretation of universally quantified subjects in sentences with negation. Like Diesing (1992), I will take this hypothesis as given.

The primary claim of the MH is that there is a correspondence between the syntactic tree and the availability of the different logical representations discussed above. Under Diesing's proposal, what she calls "tree splitting" in the syntax is analogous to partitioning a sentence into the tripartite structure of operator, restrictor and nuclear scope in the semantics. Tree splitting is induced by the presence of Q-advs and of phonologically silent operators like GEN and provides an algorithm for defining the restrictor and scope of the operator when they are present in the semantic representation of the sentence. Tree-splitting only occurs when there is an operator that requires a restriction. Otherwise, there is no partition formed in the tree, which is reflected in the semantic representation by the absence of a restrictor (and of a quantifier).

Overt Q-advs and quantifiers or other covert operators that induce the formation of a restrictive clause take sentential scope. Implicit ∃-closure binds any unbound variables in the nuclear scope. In the mapping from S-Structure to LF, material from the VP is mapped into the nuclear scope and undergoes ∃-closure while material from the IP is mapped into the restrictive clause and is bound by GEN. Diesing's (1992) division is stated explicitly in (29), and illustrated below in Figure 1.

29) *Mapping Hypothesis*

Material from VP is mapped into the nuclear scope.

Material from IP is mapped into a restrictive clause.

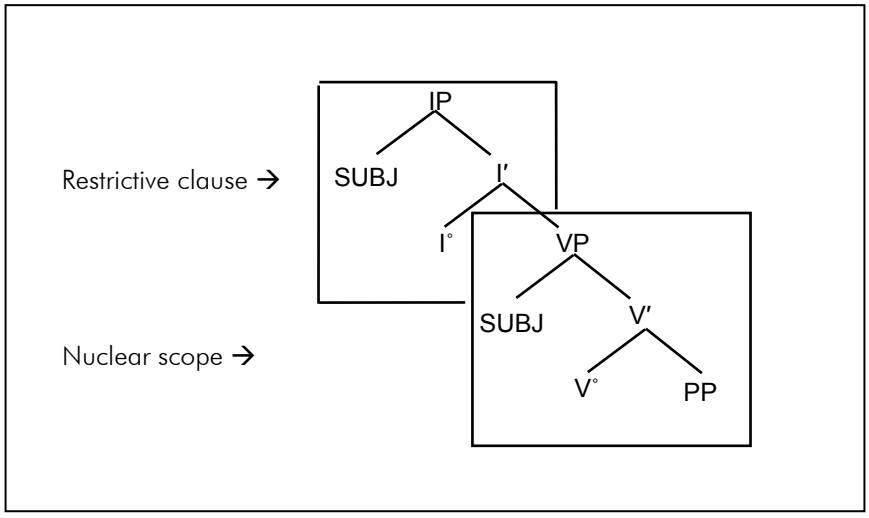


Figure 1. Tree splitting

This means that indefinites, since they are variables and derive their quantificational force from other elements in the sentence, will have different interpretations if they are interpreted in IP than they will if they are interpreted in VP[19]. Indefinites that are interpreted in IP will get their quantificational force from whatever Q-adv or covert operator has taken sentential scope and will serve as restrictors of the operator. This is what happens when an indefinite is bound by GEN. Indefinites that are interpreted in VP will get their quantificational force from implicit ∃-quantification that occurs via ∃-closure. Recall from examples like (23), repeated here, that if there is no overt operator, no restrictive clause is formed. Likewise in the syntax, tree-splitting does not occur. Since GEN behaves like overt Q-adv, the same is true for sentences without GEN – there will only be a nuclear scope in the semantic representation, provided that there is no other reason for splitting.

[19] In Diesing (1992), indefinites are not treated uniformly as variables. She differentiates indefinites that induce box splitting on their own from those that do not based on the contrast between presuppositional and cardinal determiners. This distinction is not relevant for our purposes here.

23) A mouse ate a scrap of bread.

[$\exists_{x,y}$: *x* is a mouse & *y* is a scrap of bread & *x* ate *y*]

The linguistics literature does not plainly state when a generic operator can be inserted into the syntactic (or semantic) representation. Extending Lewis (1975), Kamp (1981) and Heim (1982) to the covert generic operator, and following the discussion in Kratzer (1995) of argument structure for STAGE-LEVEL PREDICATES and INDIVIDUAL-LEVEL PREDICATES[20], I will assume that GEN may be introduced into the syntactic representation of sentences that contain stative predicates (e.g., *know French*) or habitual episodic predicates (e.g., *speak French*) when there is a variable for GEN to bind and when there is no other element in the sentence that makes the sentence true or false at a particular time or location (e.g., *yesterday*, *at the party*). GEN is not obligatorily inserted since ∃-closure is always an option (Heim 1982), but these are the conditions under which it could be part of the representation. Because the generic operator is covert, it will be useful to make this explicit by providing an example of each type of predicate with an overt Q-adv.

Following Kratzer (1995) (and Davidson 1967), individual-level (I-level) predicates, like *know French*, do not include an abstract spatiotemporal variable in their representation. Thus, overt expressions of quantification have nothing to bind when they occur with stative predicates and are ungrammatical (30a) unless there is another variable in the sentence that the adverb can bind (e.g., an indefinite NP like *French majors* in example (30b)).

[20] STAGE-LEVEL PREDICATES are, roughly, eventive predicates that denote temporary properties and are distinguished from INDIVIDUAL-LEVEL PREDICATES, which are, roughly, stative predicates that denote permanent properties.

30) a. * Judy usually knows French.

usually$_x$ [Judy knows French]

b. French majors usually know French.

usually$_x$ [[*x* is a French major] [*x* knows French]]

Eventive stage-level (S-level) predicates like *speak French* stand in contrast to stative predicates because their representation does include an abstract spatiotemporal "situational" variable, *s*, that can be bound by overt quantificational expressions. When there is another variable in a sentence, for instance, the variable introduced by an indefinite, that variable can also be bound by the operator as in the representation of (31b).

31) a. Judy usually speaks French.

usually$_s$ [[*s* is a situation that contains Judy] [Judy speaks French in *s*]]

b. French majors usually speak French.

usually$_x$ [[*x* is French major] [*x* speaks French]]

In the absence of an overt Q-adv, GEN can be introduced into the syntactic representation in the same way an overt Q-adv can. Note that in (33a), the generic operator is quantifying over situations that contain Judy, but not making a generic claim about Judy. For Diesing (1992), the generic interpretation of an indefinite depends on the indefinite falling in the restrictor of the

generic operator. The semantic representation in (33b) conveys a generic property of French majors[21].

32) French majors know French.

GEN_x [[*x* is a French major] [*x* knows French]]

33) a. Judy speaks French.

GEN_s [[*s* is a situation that contains Judy] [Judy speaks French in *s*]]

b. French majors speak French.

GEN_x [[*x* is French major] [*x* speaks French]]

Diesing (1992) motivates the MH by looking in detail at data from English and German showing that the actual interpretations of indefinite subjects correspond to different positions in the syntactic representation. The differences that Diesing proposes are relevant at LF in English, but they occur overtly at S-structure in languages like German and Dutch.

[21] There is an additional reading, although an odd one that is unlikely to be true, for each of the sentences in (31b) and (33b). If the Q-adv binds only the situation variable, the following are possible representations. The approximate interpretations for these sentences are indicated below the semantic representation.

31) b. French majors usually speak French.
usually$_s$ [[*s* is a situation] [$\exists_y$: *y* is a French major & *y* speaks French in *s*]]
In most situations, there are French majors speaking French.

33) b. French majors speak French.
GEN_s [[*s* is a situation] [$\exists_y$: *y* is a French major & *y* speaks French in *s*]]
In general, there are French majors speaking French.

These readings have been included here because with certain predicates like *available* (Diesing 1992), this alternative generic is much more plausible. See fn. 22.

German and Dutch have two overt subject positions: one inside VP and one outside of VP. In the following examples from Diesing (1992: p. 36), the difference in word order reflects the two available subject positions. The two sentence particles, *ja doch*, appear at the left periphery of the VP in German. In the sentence in (34a), the subject NP *Linguisten* "linguists" appears to the right of these particles and consequently is interpreted in the lower, VP-internal position. In this case, there is no restrictive clause because there is no operator to induce its formation. Instead, *Linguisten* has undergone ∃-closure in the VP. The sentence in (34b) provides an example of the same NP interpreted outside of VP in [Spec, IP]. Here the generic operator has induced tree-splitting. The NP *Linguisten* is the restrictor of the generic operator.

34) a. …weil ja doch Linguisten Kammermusik spielen.
since PRT PRT linguists chamber music play
"…since there are linguists playing chamber music."

b. …weil Linguisten ja doch Kammermusik spielen.
since linguists PRT PRT chamber music play
"…since (in general) linguists play chamber music."

However, not all predicates allow both of these interpretations for their indefinite subjects (Carlson 1977b). S-level and I-level predicates are distinguished along this boundary. S-level predicates, like *spielen* "play" in (34) induce both ∃- and GEN- interpretations (quantification) for

their subjects[22]. As indicated in the gloss, the sentence in (34a) states that there are linguists who are playing chamber music (at some point in time, probably now), but the sentence in (34b) indicates that it is a general property of linguists that they play chamber music. Compare the available interpretations in (34) to those in (35), also from Diesing (1992), her (37).

35) a. … weil Wildschweine ja doch intelligent sind
since guinea pigs 'indeed' intelligent are
"…since (in general) wild boars are intelligent."

b. *?… weil ja doch Wildschweine intelligent sind
since 'indeed' wild boars intelligent are

The subject of an I-level predicate can only appear in the outer, VP-external subject position. Diesing notes that for I-level predicates, a bare plural subject to the right of the particles requires a marked intonation pattern to be awkward at best and would still indicate something like the gloss in (35a). The existential reading is not possible, regardless of the intonation contour.

While this distinction is not overtly expressed with different word orders in English, we see the same interpretational constraints with S- and I-level predicates that have been provided here for German. It makes sense to say that English is the same; however, the distinction would seem to occur covertly, as Diesing proposes, at LF in English.

[22] Diesing also provides the following example where an S-level predicate induces both a generic and existential interpretation of the subject at the same time.

i. GEN_s [[*s* is a time] [$\exists_x$: *x* is a fireman & *x* is available at *s*]]
Existential generic: "It is a generic property of situations that there are firemen who are available."

Diesing (1992) begins her discussion of English by focusing on sentences like those in (36), noting that indefinite subjects can receive a generic or an existential interpretation (as observed by Carlson 1977b), even though they occur in the same position, overtly. The statement in (36a) is most naturally (and uniquely) interpreted as a statement about cantaloupes in general, rather than about any particular cantaloupes, and receives a generic interpretation. The statement in (36b), however, refers to cantaloupes that have gone bad.

36) a. Cantaloupes are unsuitable for eating.

b. Cantaloupes are rotting on the counter.

To derive these readings, Deising (1992) assumes that the bare plural NP *cantaloupes* is introduced into the restrictive clause and bound by GEN in (36a), while in (36b) the bare plural appears in the nuclear scope and is bound by ∃-closure resulting in the following two representations.

37) a. GEN_x [[x is a cantaloupe] [x is unsuitable for eating]]

b. [$\exists_x$: x is a cantaloupe & x is rotting on the counter]

As noted above for German, not all predicates allow both the generic and the existential interpretation for indefinite subjects. Examples of the distinction between S-level and I-level predicates in English are provided below in (36). As in German, S-level predicates induce both ∃- and GEN- interpretations (quantification) for their subjects. The sentence in (38b) indicates that there are activists that are available (at some point in time), but the sentence in (38c) indicates

that it is a (necessary) general property of activists that they are available. However I-level predicates, as in (39), permit only a generic interpretation.

38) a. Activists are available.

b. [$\exists_x$: *x* is an activist & *x* is available]

c. $GEN_{x,s}$ [*x* is an activist & *s* is a time] [*x* is available at *s*]

39) a. Violinists are intelligent.

b. Opera singers know Italian

Indefinite subjects of S-level predicates can appear either in the nuclear scope (to be bound by ∃-closure) or in the restrictive clause (to be bound by either an overt operator or covert GEN).

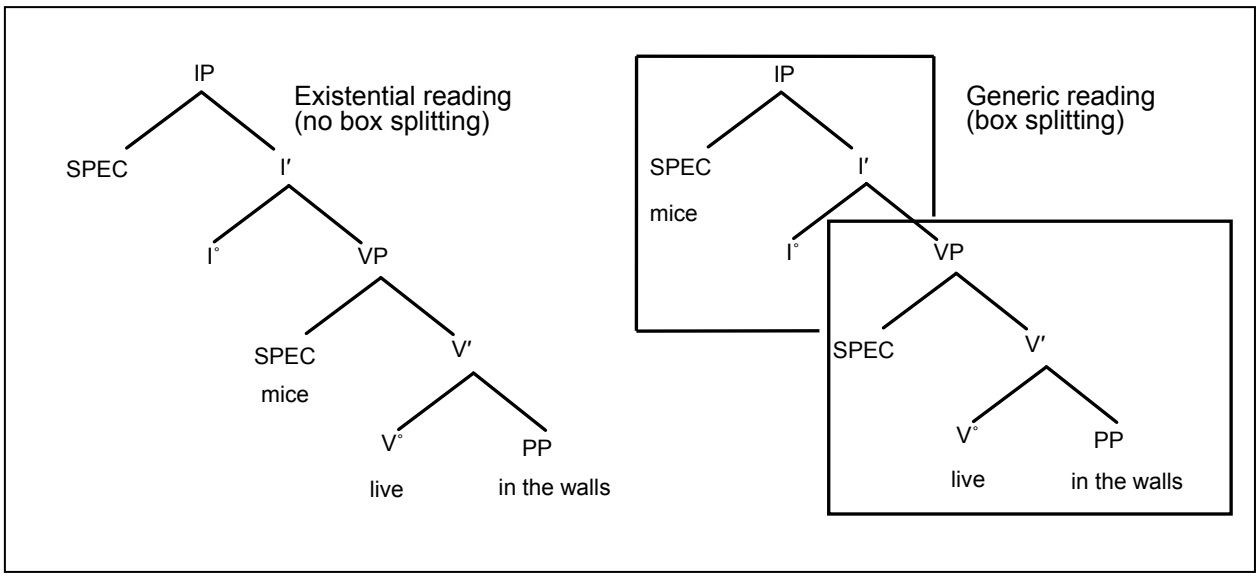


Figure 2. Syntactic representations for S-level predicates

Indefinite subjects of I-level predicates can only appear in the restrictive clause, and thus can never be interpreted existentially.

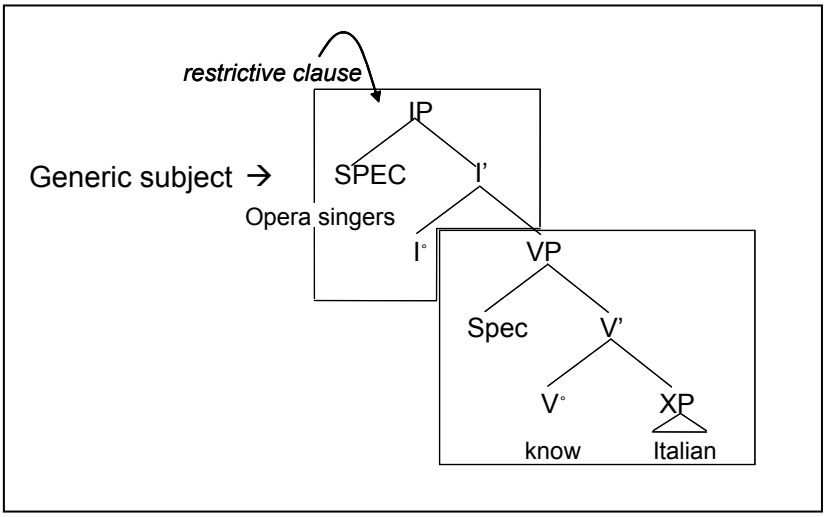


Figure 3. Syntactic representation for I-level predicates

Diesing (1992) explains the difference between these S-level and I-level predicates by positing that the former are raising predicates and the latter are control predicates. Raising predicates have an unaccusative Infl. The subject is base-generated in VP, where it is assigned a θ-role by the verb. The subjects of such predicates raise at S-structure for case but can leave a trace in [Spec, IP] and re-lower to be interpreted in [Spec, VP] at LF. Diesing points out that this analysis is similar to what we see with respect to scopal interpretations with raising predicates like *seem* in the following examples, her (13):

40) a. Firemen$_i$ seem to their$_i$ employers to be available.

b. Gila monsters$_i$ seem to their$_i$ predators to be visible.

Here, the only possible interpretation is generic because lowering would result in the subject NP's failure to c-command the co-referential pronoun. In sentences without the pronoun, both interpretations are possible, as illustrated by the following example from Diesing, her example (15):

41) a. Firemen seem to the mayor to be available.

b. Gila monsters seem to the coyotes to be visible.

I-level predicates are not raising predicates. Rather, they are control predicates with an Infl that assigns a θ-role similar to "has the property *x*", where *x* is the property expressed by the predicate (Diesing 1992: p. 26). The lexical NP in [Spec, IP] controls a PRO in [Spec, VP] (which gets its θ-role from verb). Diesing bases this analysis on the analysis of control predicates like *be anxious*, where subjects do not lower at LF since subjects are base-generated in [Spec, IP] and get their θ-role via the PRO in [Spec, VP]. Subjects of I-level predicates do not lower and always get generic interpretation[23].

42) a. $[_{IP}$ Opera singers $[_{VP}$ PRO know Italian]].

b. GEN_x [[*x* is an opera singer] [*x* knows Italian]]

[23] Diesing points out that there is potentially a problem with this analysis. PRO is supposed to be ungoverned (Chomsky 1981); however, here, PRO gets its θ-role and is governed by V. Diesing suggests that, possibly, Chomsky (1981) is wrong and PRO is governed in English (which has been suggested by a number of other researchers, see Diesing p. 26) or following Pesetsky (1989), the PRO in [Spec, VP] is governed but moves to an ungoverned, external position that is generated only when necessary. A detailed discussion of this topic is beyond the scope of this dissertation, so I will assume that Diesing's analysis can be supported and will not discuss it further.

The examples in this section have been used to illustrate Diesing's (1992) claim that there is a correspondence between the syntactic tree and the logical representations available (e.g., in (38)). The VP-region maps onto the nuclear scope and ∃-readings; the IP-region maps onto the restrictive clause and GEN-readings. Subjects of S-level predicates can appear in [Spec, IP] or [Spec, VP]; subjects of I-level predicates can appear only in [Spec, IP]. Tree-splitting occurs at LF in English, when the S-structure subject, which has moved to [Spec, IP] for case may optionally lower to [Spec, VP]. In this case the sentence gets an ∃-reading. If the subject remains in [Spec, IP] the sentence gets the GEN-reading. Diesing motivates this distinction by proposing that the two different types of predicates have different IP structures.

A consequence of tree-splitting and the resulting mapping to semantic interpretation is that indefinite direct objects (because they are in VP) should be interpreted existentially. This lines up with claims elsewhere in the literature. Note the following analysis from Krifka et al. (1995) for the interpretation of (43) and (44).

43) Unicorns have a horn.

GEN_x [[x are unicorns] [$\forall_z$ [$z \leq_a x \rightarrow$ [$\exists_y$: y is a horn & z has y]]][24]

44) Mary smokes cigarettes/a cigarette after dinner.

GEN_s[[s is after dinner & Mary in s] [$\exists_x$: x are cigarettes & Mary smokes x in s]]

GEN_s[[s is after dinner & Mary in s] [$\exists_x$: x is a cigarette & Mary smokes x in s]]

[24] Krifka et al. use this notation to specify that unicorns have one horn, rather than multiple horns (in which case they would no longer be unicorns).

The semantic representation of (43) says that if *x* are unicorns, then for every unicorn (for atomic part *z* of *x*), there is a horn (*y*) and *z* has it. In (44), the two representations differ in whether they restrict the situation to after-dinner situations that contain Mary and any number of cigarettes, as in the first semantic representation, or to after-dinner situations that contain Mary in which she smokes just one cigarette. For my purposes, what it is important to note is that the direct object NP gets an existential interpretation. It may, indeed, have a generic "feel", however this comes from its being bound by existential closure in the scope of a generically bound situation variable. Like a number of other examples that have already been discussed (e.g., (6b) and (17)), however, this too is an utterance that makes generalizations about situations that contain Mary and cigarettes, not about cigarettes themselves.

One class of expressions that are an exception to the generalization that indefinite direct objects are interpreted existentially is PSYCH VERBS[25]. These are verbs like *appreciate*, *fear*, *frighten*, *love*, *like*, *hate* (see Levin 1993 for the full list of such examples) that are considered to be I-level predicates. As such, they require the generic interpretation of their indefinite subjects (as detailed above); however, these verbs also require a generic interpretation of their direct objects. This is true for English, as well as German and Dutch. The analysis in the literature (discussed in Diesing 1992) is that direct objects of this class of verbs undergo scrambling at LF to adjoin to IP, mapping into the restrictor by the tree-splitting algorithm in the same way subjects do.

It is important to acknowledge that the classification of predicates as S-level or I-level is sometimes problematic. There are a number of examples where context plays an important role.

[25] Diesing does not distinguish between subject experiencer (e.g., *love*) and object experiencer (e.g., *frighten*) psych verbs; however, this generalization seems to be true of both classes.

For instance, emotional states that make up the class of EXPERIENCER PREDICATES including *angry*, *cheerful*, *obnoxious* and *nervous* might seem like S-level predicates, but in fact they behave in some ways like I-level predicates. Note that *angry* is not acceptable in *there*-insertion contexts.

45) a. Insurgents are angry.

b. * There are insurgents angry.

c. ? There are barbarians angry (at the gates).

However, their unacceptability is not as marked as with canonical I-level predicates like *intelligent*. Compare the following to the examples in (45).

46) a. Cephalopods are intelligent.

b. * There are cephalopods intelligent.

Moreover, they can occur with progressive-*be* in English, which clearly gives rise to S-level interpretations. In such cases these predicates allow *there*-insertion:

47) a. There are hobbits being cheerful.

b. There are Chihuahuas being nervous.

Diesing (1992) suggests that progressive-*be* selects an S-level Infl in contexts where its meaning corresponds to the verb *act* and the adjectives have an adverbial (not predicative)

function. I-level predicates that cannot modify *act* are not acceptable here (48b). Such (adverbial) interpretations of the predicate are clearly distinguished from true S-level transitory properties, which cannot occur with progressive be (48c):

48) a. Marija is being intelligent/cheerful/nasty/nervous.

b. * James is being tall.

c. * Firemen are being available.

Contextual effects related to the status of a predicate as S- or I-level affect predicates other than just the experiencer predicates discussed above. Note the following examples from Diesing (in her examples (68) – (71)) where the interpretation of canonical stage level predicates like *be available* and *be sick* are subject to context:

49) a. There are children sick.

b. Children are sick. (ambiguous)

c. * There are children with red rashes sick.

d. There are children with red rashes available.

e. Children with red rashes are sick/available. (generic only?)

Although it is not always trivial to classify a verb as S- or I-level, there are still broad generalizations that can be made about these two categories. The categories are cohesive, with these few exceptions that form the relatively homogenous class of expressions related to psychological states.

Diesing (1992) provides additional support for the MH with data from English and especially from German, where many of the relevant contrasts occur overtly at S-structure. However, a more detailed description of the intricacies of Diesing's work is beyond the scope of this dissertation. The central idea is that the overt syntactic distinctions in German that result in differences in interpretation are the same processes that occur in English. The only difference is that in English, these things happen covertly, at LF[26]. I use Diesing's Mapping Hypothesis for concreteness, but it is important to understand that the precise mechanism of indefinite interpretation is less important than the basic insight that lies behind this hypothesis, namely, that indefinites that occur low in the structure are assigned an existential reading whereas indefinites that occur high in the structure are assigned a generic interpretation. See Rooth 1995 and Cohen & Erteshik-Shir 2002 for alternative proposals capturing the same basic insight.

For the purposes of this dissertation, let us assume that the MH is sufficiently motivated and provides a coherent syntactic explanation of the different semantic interpretations available for indefinite NPs.

1.4 Genericity: a property of NPs or utterances?

The previous research on the acquisition of generics has contributed many interesting and valuable findings, in particular, findings that inform research on psychological essentialism and how children create categories. However, this research meets with certain explanatory difficulties when it attempts to describe the acquisition of generic language. The primary reason for this is that the extant acquisition literature focuses on the role of the NP in the genericity of

[26] The rest of Diesing (1992) motivates the MH with data on quantification including strong and weak determiners and their behavior with respect to scopal interpretations. Diesing refines her claims about the classification of indefinites based on their more general behavior with respect to QR. She also examines the influence of different semantic classes of verbs (e.g., verbs of creation and destruction) on the interpretation of indefinites.

an utterance. While it is clear from the discussion in §1.2 that the NP *can be* an important source of genericity, an NP's status as kind-denoting is relevant only in certain environments; such linguistic contexts are not part of child-directed speech[27].

To be clear, the previous acquisition literature does not explicitly say that the NP is the source of genericity. Rather, it is evident that the focus is on the NP for two main reasons. First, in taking inventory of parental uses of the different morphological forms of NPs and in examining children's interpretations of these NP forms, this research does not examine the fuller linguistic context (e.g., predicate type). Second, this literature does not distinguish genericity via a kind-NP from genericity via a characterizing sentence, which, as I have demonstrated in this chapter, are two very different linguistic phenomena. In fact, the acquisition literature never considers the possibility that the examples of generic utterances it is concerned with are best described as characterizing sentences. Instead, despite the claim in the linguistics literature that indefinite singular NPs are not kind-NPs (except on their taxonomic reading) and are interpreted generically only when they occur in characterizing utterances (see discussion in §1.2), the previous literature on the acquisition of generics wrongly treats all NPs (including indefinite singular NPs) as "kind-referring expressions":

> "…in spontaneous language, mothers likewise are more apt to discuss animal categories using 'kind-referring expressions' (Krifka et al. 1995). This result suggests that generic NPs may provide a means of enriching children's category knowledge in domains that are particularly central for adults."
>
> Gelman & Tardif (1998: p. 241)

[27] See discussion in §1.2.

However, for Krifka et al. (1995), "kind-referring expressions" are kind-NPs, a category that the reader will recall refers only to the kind as a whole and does not include indefinite singular NPs. Further, the previous research conflates bare plural NPs (e.g., *bears*) and overtly existentially quantified NPs (e.g., *some bears*).

In her discussion of the semantic features of generics, Gelman (2003) points out the fact that one feature that distinguishes generics from indefinites is that the former are generically true. She goes on to say that the difference between these two readings is "particularly critical because the same form of the noun phrase can be used for both" (p. 207) providing the two utterances in (50).

50) a. Bears live in caves.

b. I saw some bears in the cave.

While it is true that the bare plural NP and the overtly existentially quantified NP both contain the same string (e.g., *bears*), they are not the same. The existentially quantified NP contains overt quantificational information that the bare plural does not. This is exactly the kind of information a child can use to learn about generic and referring language.

These features of the previous analysis lead to the description of a system that would likely be impossible to acquire by analysis of the linguistic signal because it would require a child to derive a system of generics based on a many-to-many mapping from NP form to NP function. In fact, this is exactly what research by Gelman and colleagues has shown (Gelman 2004; inter alia): because parents use both available morphological forms of NPs in referring to generic concepts, "…what we find, through a series of converging studies, is that parental input

retains the complexity of mapping, thereby not greatly simplifying the inductive task that children face" (Gelman 2004: p. 456). In short, these authors have argued that the input by itself, except in the broadest sense (i.e., children can attend to the use of the plural when their caregiver says, for example, *They have four stomachs* in the presence of a single cow) is uninformative regarding generic interpretation. And, if the focus is on the NP, they're right – the input is not very informative. When a child encountered an NP, he would have to compute a very fine-grained analysis on individual sentences including analysis of weak cues such as "…at the very least, morphosyntactic cues, contextual cues and world knowledge … none [of which] is individually sufficient" (Gelman 2003: p. 215) to determine whether the NP should receive a generic interpretation.

One solution to this problem is to say that somehow children can make use of matrix of weak cues that the input provides. This seems unlikely, however, given the complexity of the cues. Another solution, and the one that Gelman (in Hollander et al. 2002, Gelman 2003 and Gelman & Raman 2003) adopts, is to say that instead of using such a complicated array of information to determine whether or not an NP is generic, generic interpretations are the default for children, at least by the time they are three years old. Children look for information in an utterance to tell themselves that an NP is not going to be interpreted generically and until they find such cues, they assume genericity. I provide evidence below and in Chapter 3 against this hypothesis. Instead, the solution I propose approaches the problem in a different way. Genericity is a property of the sentence grammar, not of the NP, so being generic or existential depends on the position that a constituent is interpreted in in the syntax (at LF in English, at S-structure in German or Dutch). If that's the case, then a child is not trying to use these cues to determine whether or not an NP is generic. The child's main focus on the input is to determine which

morphological forms of determiners map onto the different semantic forms. The learner can make use of some of these cues to determine whether or not the covert generic operator is present.

1.4.1 An NP-centric view of the acquisition of genericity

Word learning is a complex task on its own. While children have been shown to fast-map a label to a new word (Carey & Bartlett 1978, Heibeck & Markman 1987), there is more involved in learning the broader meaning of, for example, *dog* than simply pointing to a dog and labeling it when the animal is currently in the child's view. The string *dog* can denote a kind, when it's an NP it can denote a property, or the NP can combine with a determiner to denote an individual or a generalized quantifier (e.g., *every dog*). So children cannot simply be learning the meaning of *dog* when their parents utter the string in the presence of the animal because the intended meaning of the string depends on how it combines with the other elements in the utterance.

One part of the word learning problem is that words are not uttered in isolation; rather, they are uttered in sentences. The same string *can* be used to refer to different things depending on the type of sentence it's used in because different types (listed above) can be indicated by the same string. Considering the examples from (3), first presented in §1.0, (3a) refers to a particular barking dog, (3b) refers either to the dog that is barking, or to a non-specified dog and (3c) refers to a generic property of the kind [dog].

3) a. The dog is barking.

b. A dog is barking.

c. Dogs bark.

While some element of the meaning of *dog* may be constant across these examples, what's being referred to is determined by the whole NP (e.g., *the dog*, *a dog*, *dogs*) at a minimum or possibly by the whole sentence. As noted in §1.0, it is difficult to determine an NP's intended interpretation because a given NP form is often compatible with both generic and existential or referential interpretations depending on the sentence it occurs in and there is no unique NP form that indicates genericity, as shown in (4) and (5), repeated here.

4) a. The cow has four stomachs.

b. A camel spits three times before sleeping.

c. Tasmanian tigers are extinct.

5) a. The cow is eating.

b. A camel spit three times.

c. There are Tasmanian tigers in the natural history museum.

Because there is no one-to-one mapping between NP form and function with respect to generic and nongeneric language, if the child did not have the MH to rely on, she would have to rely on an array of weak cues that are in the input.

One important feature would be the morphosyntactic information available in an utterance. This includes determiners and number, jointly, as well as tense and aspect. Gelman (2003) points to the overlap between the formal properties of the NPs in (4) and (5), noting that generics can be expressed with the definite singular, indefinite singular or bare plural. Therefore, the speaker must also use number in combination with the determiner to decide whether

something is generic or nongeneric. Gelman provides the following example (2003: p. 215) showing that a definite determiner cannot combine with a plural NP in a generic statement.

51) a. The bears are huge.

b. Bears are huge.

Although Gelman does not explicitly say this, it should be noted that the interaction of the definite determiner and plurality can indicate a nongeneric, but there is no equivalent unequivocal indication of a generic. Instead, a child noting the interaction of the determiner and number is left with "nongeneric" or "not nongeneric".

This research also acknowledges the importance of tense, saying that "with the exception of the historic past (e.g., "*Woolly mammoths* roamed the earth many years ago") past tense utterances are not generic…" (Gelman 2003: p. 216). However, this is simplifying the issue. As Krifka et al. (1995) point out, any verbal predicate in the simple present, past tense or future tense *can* have a characterizing or particularizing interpretation. This includes, as Gelman mentions, the historic past, but it also includes the possibility of a particularizing interpretation for the simple present used in the 'reportive present'. All of these cases are non-standard uses of tense, but it should be noted that any utterance could have a characterizing or particularizing interpretation.

52) John smokes/smoked/will smoke a pipe.

Lastly, as far as morphosyntactic cues are concerned, Gelman (2003) discusses the importance of aspect as a cue because "…a statement in the simple present, such as "Cats meow," is generic, whereas a statement in the present progressive "Cats are meowing," is nongeneric…" (p. 216). While there is certainly a strong correlation between progressive and perfect sentences receiving a particularizing interpretation, Gelman simplifies the issue by neglecting to discuss that the addition of an adverb can override the present progressive's tendency toward such an interpretation.

53) Bears are usually/always/often roaming around in the woods.

Use of this type of morphosyntactic information to determine that an NP is generic would require a fine-grained analysis of each sentence. A child would have to decide what information he receives from the interaction of the determiner and plurality. And this, in itself is not an easy task because there is no direct indication of genericity of the NP: recall that for Gelman, (2003) the interaction of the definite determiner and plurality can indicate a nongeneric, but there is no equivalent direct indication of a generic. Then, the child would have to look at the interaction of the determiner and verb. Sometimes the verb tense/aspect will determine genericity or nongenericity, other times, it would be the noun. Precisely how a child would know whether to attend to the NP or VP remains unspecified. Furthermore, Gelman does not discuss the particular pieces of tense and aspectual information that she expects the child to be able to use in determining whether an utterance is generic or nongeneric. In the end, even the morphosyntactic analysis Gelman proposes is not sufficient to determine whether an NP is generic.

Children would also have to turn to context. This includes both mismatches in anaphoric relations and the sentence form. For instance, when an object is labeled with a singular NP and subsequently described with a plural anaphor, this is an indication that the speaker is using generic language. As discussed above, some of the previous research (Gelman & Raman 2003) refers to this type of information as "pragmatic context".

54) a. This is a tapir. They like to eat leaves.

b. These are my tapirs. They like to eat leaves.

In (54a) the mismatch indicates that the speaker has switched from labeling an individual member of the kind to talking about the kind as a whole. However, while the anaphoric mismatch in (54a) is a good cue to genericity, Gelman (2003) fails to point out that in (54b), the speaker may be taking about the dietary preferences of these particular tapirs *or* switching to generic talk about the kind.

The form that a sentence takes can also provide some insight into genericity. However, if children are trying to figure out which NPs are generic, this cue actually presents an even bigger problem for them because the information provided by the form of the sentence is in fact ambiguous. Consider the following two statements:

55) a. Dingoes live in Australia.

b. There are dingoes in Australia.

For Gelman (2003), there is a clear distinction between the utterances in (55a) and (55b). The utterance in (55a) is a generic statement about dingoes and the utterance in (55b) is not. However, sentences like (55) have long been noted in the linguistics literature to be ambiguous between a generic and existential interpretation of the subject. Note the following from Milsark (1974).

56) Typhoons arise in this part of the pacific.

a. Typhoons in general have a common origin in this part of the Pacific
(i.e., for typhoons it holds: they arise in this part of the Pacific).

b. There arise typhoons in this part of the Pacific
(i.e., for this part of the Pacific it holds: there arise typhoons).

As should be clear to the reader from the discussion in §1.3, the different interpretations of (56) depend on whether *typhoons* is interpreted in the restrictor and *in this part of the pacific* is interpreted in the nuclear scope (as in (56a)) or vice versa (as in (56b)[28]. Applying this to the example in (55a), the following two interpretations are available.

55) a. Dingoes live in Australia.

a'. For dingoes it holds: they live in Australia.
GEN_x [[*x* is a dingo] [*x* lives in Australia]]

[28] Although it is not always a reliable cue, pitch accent placement on either *typhoons* or *in this part of the Pacific* can determine the partitioning of the sentence into restrictor and nuclear scope. See fn. 14 in §1.3.1 for more discussion.

a". For Australia it holds: there are dingoes that live there.

GEN_s [[*s* is in Australia] [$\exists_x$: *x* is a dingo & *x* lives in Australia in *s*]]

The sentence in (55) is not unambiguously a generic statement about dingoes, so if a child is using sentence form to determine which NPs are generic, he may come to the wrong conclusion with respect to bare plural subjects.

1.4.2 A structural view makes better use of cues

Learning about genericity by focusing on the NP requires the learner to keep track of a lot of piecemeal information. Recall that a child has to keep track of morphosyntactic cues like determiner and plurality, which at best can only indicate a nongeneric. Further, the child must keep track of the interaction of the determiner and verb, which presumably leads to different conclusions that Gelman (2003) has not specified. As if this were not complicated enough, a child then has to turn to sentence form, which is not, in fact, a clear indicator of genericity.

While not impossible, it seems doubtful that the language learner could use the complex matrix of cues described here to determine whether an NP is generic. This is precisely why Hollander et al. (2002), Gelman (2003) and Gelman & Raman (2003) propose that generic interpretations are the default for children and that in the absence of specifying information, children will assume that an utterance is generic. However, this proposal too is unlikely. Consider again examples like (55). This sentence lacks all the specifying information that could reveal its nature as a referring sentence. But as I have shown, (55) is ambiguous. I do not wish to claim that the cues to genericity that Gelman discusses are not at all useful. If, instead of focusing on the NP, we focus on syntactic structure as an indicator of genericity, the types of

cues that I have just reviewed are precisely the kinds of cues that can be useful. These are the kinds of cues that can tell a child whose grammar includes the MH that the generic operator is licensed.

Recall the discussion in §1.3.3 about when GEN can be introduced into the syntactic representation (Davidson 1967, Kratzer 1995). GEN can be introduced into the representation for sentences that contain I-level predicates (e.g., *know French*) or S-level predicates (e.g., *speak French*) when there is a variable for GEN to bind and when there is no other element in the sentence that makes the sentence true or false at a particular time or location (e.g., *yesterday*). Let us examine how the cues in the input can be used to determine this.

For instance, *be huge* is an I-level predicate. When the subject is a definite NP, there is no variable for GEN to bind, so it is not present in the representation. However, an indefinite NP like the bare plural *bears* introduces a variable to the semantic representation, so the generic operator is present because there is no other information in the sentence that makes the sentence true or false at a particular time. *Smokes* is an S-level predicate with an abstract spatiotemporal argument for GEN to bind, and there is no other information in a sentence like *John smokes a pipe* to indicate that the sentence is true or false at a particular time. Thus, it is likely that GEN would be introduced into the interpretation. We might predict that a child is likely to misinterpret instances of the reportive present for this reason and would have to learn about this option later. The same is true for the uses of past and future tense and the progressive aspect in characterizing sentences, rather than in their canonical particularizing forms.

Further, if a child's grammar includes the MH, she will know automatically that both interpretations are available for utterances like (55). Consider the following:

57) a. A: What do you know about Dingoes?

B: Dingoes live in Australia.

GEN_x [[x is a dingo] [x lives in Australia]]

b. A: What do you know about Australia?

B: Dingoes live in Australia.

GEN_s [[s is in Australia] [$\exists_x$: x is a dingo & x lives in Australia in s]]

The child can make use of the prior discourse context to determine whether the generic operator is indeed present, or if *dingoes* has been bound by ∃ through ∃-closure.

1.5 Learning algorithm

Using Diesing's (1992) MH as a starting point, I propose a learning algorithm by which a child could exploit the input to learn about the acquisition of generics. In order to do this, I assume the following: the MH is part of UG, the inventory of determiner meanings in language is made available by UG and children know something about the relation between novelty and familiarity on the one hand and indefinite- and definiteness on the other.

The MH provides a clear, linguistically well-motivated proposal about the interpretation of indefinites that accounts for the different interpretations that are available in the adult grammar for bare plural and indefinite singular NPs: syntactic position at LF determines interpretation. The MH is something that could act as a filter on the input, informing acquisition by constraining what children take from caregiver speech. Let us assume that UG provides this mapping to children and that children come to the task of learning about generic and referring language knowing that quantificational operators (overt and covert), when present, induce tree-

splitting in the syntax, which in turn results in the formation of a restrictive clause and a nuclear scope in the semantics. Whatever is interpreted high in the tree is interpreted within the restrictive clause and whatever is interpreted low in the tree is interpreted within the nuclear scope. With respect to the presence of the covert generic operator, children know that indefinites that occur in the restrictor are assigned a generic interpretation whereas indefinites that occur in the nuclear scope are assigned an existential reading. If there is no quantificational operator (overt or covert), there is no tree-splitting, so in the semantic analysis there is only a nuclear scope in which the variables are bound by ∃-closure.

If UG provides children with the MH, then making use of it to determine what is generic and what is existential requires knowing what morphological determiners map onto the semantic category of indefinite determiners. Let us also assume that there is a universal inventory of determiner meanings made available by UG (May 1991; Partee 1992; Chierchia 1995, 1998; Matthewson 2004, cited in von Fintel & Matthewson (in press))[29]. For instance, some determiners are "Kamp-Heim indefinites" and combine with an NP to produce expressions that introduce a variable to the semantic representation. Other determiners in the inventory are "definite". These combine with an NP to produce an expression that denotes a unique individual. Still other determiners are universal quantifiers. They combine with an NP to produce an expression that denotes a set of all the members of a set. The list of semantic determiners provided here is not exhaustive. If UG provides an inventory of possible determiner meanings, then all the child has to do is determine the mapping between the semantic determiners (e.g.,

[29] See von Fintel & Matthewson (in press) for a review of the current research on semantic universals for the lexicon, functional morphology and pragmatics.

"Kamp-Heim indefinite", "ι", "∀") and the morphological determiners (e.g., *a*/*an*, *the*, Ø) in his language.

I propose that children are able to determine the mapping between morphological and semantic determiners by attending to the fact that certain morphological determiners are used with an NP when it is mentioned for the first time in a discourse and that other morphological determiners are used when the NP is mentioned the second and subsequent times in the discourse. Their knowledge of these distinctions falls out of the assumption that the inventory of determiner meanings is specified by UG. If being a "Kamp-Heim indefinite" means that something introduces a discourse referent, and if this is specified in UG, the learner can make the inference from seeing an NP with *a*/*an* as the determiner being used to introduce a discourse referent to the knowledge that NPs with *a*/*an* as the determiner have Kamp-Heim semantics (i.e., they introduce a discourse referent). Further, children can attend to where in the sentence these different morphological forms occur. For instance, subjects are more likely to represent information that has already been mentioned in the discourse, and as a consequence, are more likely to be definite than are non-subjects (Prince 1992)[30]. New information, canonically introduced by indefinites, is more natural following given information and is more common later in the sentence. Additionally, Fisher & Tokura (1995) found that in a story-retelling task in which mothers watched a series of mini-plays and then related the story either to their 14-month-old child or to the experimenter, mothers were "reluctant" to use the definite article in child-directed speech. Repeated words were more likely to be definite in speech directed to both groups (i.e., an indefinite article was only rarely used for a repeated target); however mothers

[30] Prince's analysis considered only "canonical," unmarked subjects (i.e. in English, subjects that occur in a pre-verbal position).

were significantly less likely to use the definite with children than with adults[31]. This suggests another difference between definite and indefinite NPs that children could use to distinguish the two: overall usage. Children therefore know something about novelty and familiarity and indefinite- and definiteness by virtue of the fact that indefinites and definites exhibit different patterns of usage. If children are able to distinguish indefinites from definites in this way, they know that only the morphological forms that correspond to indefinites matter for the MH and can infer, further, that both existential and generic readings are licensed under appropriate syntactic and semantic conditions.

To illustrate how the proposed learning algorithm would work, consider the situation when the statements in example (58) are uttered in sequence in an out-of-the-blue context.

58) a. I see a bagel.

b. The bagel has sesame seeds on it.

When a child hears these sentences, how does he decide what they mean? How does he know whether they are characterizing or particularizing? According to the learning algorithm, the child would hear the string *bagel* in (58a) and note that this is the first time the string has been used in the discourse[32]. The NP is produced with a particular morphological determiner (*a*) and it occurs in direct object position in the sentence. These facts provide the child with an example of an NP (*bagel*) with *a* as the determiner being used to introduce a discourse referent. The child thus has evidence that *a* lines up with the semantic indefinite determiner. Further, the child is provided

[31] The authors suggest that mothers thought that even visible referents were less uniquely identifiable to children than they were to adults, making them less ideal for use of the definite.
[32] Recall that for the purposes of this example, this is an out-of-the-blue context.

with a piece of evidence about the distribution in the syntax of old and new information. Then, the child hears the string *bagel* in (58b) and notes that the second time the string *bagel* was been mentioned, it was produced with *the* and it occurred in subject position. This provides the child with an example of an NP with *the* as the determiner being used to refer back to a discourse referent thus providing the child with evidence that *the* lines up with the semantic definite determiner (ι). The sentence in (58b) provides the child with another piece of evidence about the distribution in the syntax of old and new information. The child should also note, upon encountering (58b) that this is the first time the string *sesame seeds* has been used in the discourse. This NP is produced as a plural without an overt determiner (a bare plural) and it occurs as the direct object of the sentence. As with *a bagel*, in (58a), these facts provide the child with an example of a particular morphological form of NP (bare plural) being used to introduce a discourse referent (*sesame seeds*). The child can make use of this information to map bare plural NPs onto the semantic indefinite determiner. This provides the child with another piece of evidence about the distribution in the syntax of old and new information. My learning algorithm assumes that children are able to make a distinction between indefinites and definites on the basis of information in the input similar to what I have described here.

Once children have been able to distinguish definites from indefinites, the MH takes over. Because their grammars include the MH, children know that sentences with indefinite NP subjects, unlike those with definite NP subjects, are ambiguous between receiving a generic and an existential interpretation. Recall from the discussion in § 1.3.3 that the MH tells children that sentences with indefinites that are interpreted outside VP at LF are assigned a generic interpretation whereas those with indefinites that occur within VP are bound by ∃-closure and assigned an existential reading. If there is no quantificational operator (overt or covert), there is

no tree-splitting in the syntax, so in the semantic analysis there is only a nuclear scope in which the variables are bound by ∃-closure. Recall also that the generic operator is licensed when the syntactic representation of a sentence contains stative predicates (e.g., *know French*) or habitual episodic predicates (e.g., *speak French*) that have a variable for GEN to bind and when there is no other element in the sentence that makes the sentence true or false at a particular time or location (e.g., *yesterday*, *at her house*), but that GEN need not be inserted even when it is licensed because ∃-closure is always an option.

The sentence in (58a) contains an episodic predicate (*see*) and no spatiotemporal information that would prevent GEN from being inserted into the representation. Since *see* is an S-level predicate, such a sentence could be interpreted generically (if it had an indefinite subject that was interpreted in IP) or existentially (if it had an indefinite subject interpreted in VP); however, the only variable in the sentence was introduced by the direct object, which occurs in VP. Material inside the VP is bound by ∃-closure, so the child interprets *a bagel* existentially. The generic operator is not introduced into the representation because there is no variable for it to bind. The sentence in (58b) contains a stative predicate (*have*) and no spatiotemporal information that would prevent GEN from being inserted into the representation. Since the MH is not relevant for definite NPs, my learning algorithm predicts that the child would overlook *the bagel*, which as a definite NP does not introduce a variable for GEN to bind. In this example, too, the only variable in the sentence is introduced by the direct object, which the child would interpret existentially. The generic operator is not introduced into the representation because there is no variable for it to bind.

Let us now imagine that following (58b) the child heard the utterance in (58c).

58) a. I see a bagel.

b. The bagel has sesame seeds on it.

c. A bagel has a hole in the middle.

The sentence in (58c) contains a stative I-level predicate (*have*) and like the other two sentences, there is no spatiotemporal information that would prevent GEN from being inserted into the representation. In this case, there are two variables present in the representation, the variable introduced by *a bagel* and the variable introduced by *a hole*. This time GEN is introduced into the representation because there is a variable for it to bind. Because the MH is part of the child's grammar, the child knows that *a bagel*, is interpreted outside of VP in the syntactic structure, receives a generic interpretation, whereas the direct object *a hole*, which occurs in VP. Material inside the VP is bound by ∃-closure, so the child interprets *a hole* existentially.

What sets my learning algorithm apart from previous accounts is that I claim that genericity is best understood as a property of sentence grammar, not as a property of NPs. Under this learning algorithm, children do not need to use the input to figure out which NPs receive generic interpretations and which NPs receive existential interpretations. They know this information by virtue of the MH being part of their grammars. Children need the input, but to determine which morphological forms of NPs are indefinites, definites, or quantificational NPs.

1.6 What do children know?

In §1.4, I have established what the pieces of the grammar of genericity are. In §1.5, I have sketched a learning algorithm and laid out my assumptions about the kinds of knowledge that UG would have to provide to children in order for them to make use of my learning algorithm. If

my assumptions and proposed learning algorithm are reasonable, we should expect to find evidence supporting my claims and proposal in the literature. More specifically, we should expect to find evidence that children are aware of the two subject positions and that they are aware of the semantic consequences of the two positions. We should expect that there is enough information in the input to children that they can determine which morphological determiners in their language map onto the inventory of semantic determiners provided by UG. We should also expect to see evidence that children can distinguish definite and QNPs from indefinites.

1.6.1 Evidence for two subject positions

Recall from §1.3.3 that a crucial part of the motivation for the MH involved the distinction between S- and I-level predicates. Recall also that the subjects of I-level predicates can only receive a generic interpretation, while subjects of S-level predicates are ambiguous between a generic and an existential interpretation depending on where in the clausal architecture they are interpreted. If children are sensitive to this distinction, we would have indirect evidence that children are aware of the two subject positions postulated in the MH. While the distinction between S-level and I-level predicates is one that might, at first, seem beyond the knowledge of a preschool language learner, this is not the case. There is ample evidence to suggest that children would be able to profit from the distinction between these two predicate types in regard to subject interpretation. In fact, children as young as 2-years-old have been shown to be sensitive to the I-level/S-level distinction.

Becker (2000) found that children are aware of the distinction between S- and I-level predicates based on their use of the copula. Children systematically drop the copula for S-level predicates, but produce an overt copula for I-level predicates. Research by Graham and

colleagues (Graham et al. 2003, Graham et al. 2005) provides further support for this claim because it shows children treat I-level predicates like *friendly* different from S-level predicates like *hungry*. Graham et al. 2003 examined whether 4-year-olds' willingness to extend three different types of familiar adjectives was dependent on the type of trait that the adjective expressed. Graham et al. did not refer to these adjectives as S- or I-level predicates, however, the types of adjectives they used line up relatively closely with these categories. Two of Graham et al.'s categories, "transient emotional states" (e.g., *happy*), and "transient physiological states" (e.g., *hungry*), map onto S-level predicates while the other, "stable traits" (e.g., *gentle*), was comprised of I-level predicates. Graham et al. found that children treated these two types of adjectives differently. They extended "stable traits" significantly more often to other exemplars of the same kind than they did "transient emotional states" or "transient psychological states".

If children know about the distinction between S-level and I-level predicates, it is also reasonable to assume that they know something about the structure of the different types of Infls that go along with each of these predicate types (i.e., a raising-type Infl for S-level predicates and a control-type Infl for I-level predicates). Children should know that the indefinite subjects of I-level predicates are base generated in [Spec, IP], that they must be interpreted there and that they always receive a generic interpretation simply as a consequence of how the syntax maps onto the semantics of an utterance with an I-level predicate. We can assume that children's familiarity with S-level predicates would include the knowledge that subjects of S-level predicates are base-generated in [Spec, VP]. We can also assume that children know about subjects raising to get case, or for interpretation as this is motivated for other types of sentences elsewhere in the linguistics literature.

What children might not know is that subjects of S-level predicates can be interpreted in VP, even after raising to [Spec, IP] for case, so they may not entertain the possibility that subjects can be interpreted existentially, like other NPs that are interpreted within VP (i.e., direct objects). Previous research does not bear on this question. However, it will be addressed in Chapter 3.

1.6.2 Distribution of generics and nongenerics in child-directed speech

Another important part of the assumptions I make is that there is enough information in the input to children that they can determine which morphological determiners in their language map onto the inventory of semantic determiners provided by UG. A number of studies by Susan Gelman and colleagues (Gelman 2004, inter alia) bear on this issue.

Pappas & Gelman (1998) examined whether the form of NP that was used differed in sentences that were interpreted generically and nongenerically in child-directed and child speech. Specifically, they investigated the use of singular or plural NPs to describe the presentation of one or multiple exemplars of a kind, and the interaction between the two. By their classification, indefinite singular (e.g., *a crab*) and bare plural NPs (e.g., *crabs*) could be generic expressions. Plural indefinites (e.g., *some/two/many crabs*) and definite NPs (singular and plural) were always classified as nongeneric. They examined 26 mother-child pairs (n = 12 aged 1;11 – 3;0 and n = 14 aged 3;2 – 4;9[33]) and videotaped the pair going through a picture book that varied whether there was a single instance (e.g., one crab) or multiple instances (e.g., many crabs) of a kind of animal. Animals were chosen because previous research had shown that parents engage in more generic talk with animals than with artifacts (Gelman 1988, Keil 1992).

[33] Children's ages are presented in the following format: year;month.day.

Overall, they found that both adults and children used more nongeneric utterances than generic utterances. However, 24 of the 26 mothers produced at least one generic with overall rates across mothers ranging from 0% - 41% of the utterances the mothers produced. They also found that nongeneric language was used more often to talk about the kinds that appeared on multiple-instance pages, and that generic language was used more often to talk about the kinds that appeared on single-instance pages. Furthermore, the form of nongeneric NPs was closely linked to the structure of the page (i.e., singular NPs were used more often when a single instance was presented; plural NPs were used more often when multiple instances were presented), but for generics, the morphological form of the NP was independent of the information depicted. In fact, plural NPs were as frequent when only one instance was presented as when multiple instances were presented.

These findings are relevant for the literature on conceptual development, but they are also relevant in a purely linguistic sense because they show that parents are using generic expressions in their speech to children, and that they use them in different environments from nongenerics. Pappas & Gelman (1998) do not report on the overall distributions of the particular morphological forms of NPs, but the fact that parents make a distinction in their expression of number between generic and nongeneric utterances reflects the different status that the two types of expressions have. The fact that parents are using both forms indicates that children are getting evidence for both of the morphological forms of determiners that map onto the semantic indefinite determiner. This suggests that the input to children could indeed give them insight into the mapping between morphological determiners and semantic determiners because it is sufficiently rich and reflects the complexity of the mapping.

1.6.3 Cross-linguistic comparison of generics in child-directed speech

In a series of experiments, Gelman & Tardif (1998) investigated the use of generic expressions cross-linguistically, comparing how generics are used and interpreted in English and Mandarin Chinese. The comparison of English and Chinese was motivated by their view that the NP is the most important clue to genericity. Chinese does not require the use of articles and the formal marking of generics is less obligatory and less overt in Mandarin than in English. Note the following example from Gelman & Tardif showing that the same string in Mandarin reflects three different English glosses.

59)	xiao3	ya1zi	yao2yao2bai3bai3		de	zou3	lu4
	little	duck	waddlingly		DE	walk	road

a. The duck is waddling.

b. The ducks are waddling.

c. Ducks waddle. / A duck waddles.

So while the marking of genericity is not always obligatory or overt in English and different NP forms can be used in generic statements, English does at least require the use of articles for most nouns (bare plural and mass nouns being the two exceptions). Within each language, Gelman & Tardif also wanted to look at which semantic domains received the most generic talk.

In their first experiment, Gelman & Tardif (1998) examined child-directed speech in caregiver-child pairs (Mandarin, n = 10 aged 1;9 – 1;11, mean MLU 1.82; English, n = 10 aged 1;9 – 1;11, mean MLU 1.30 and n = 10 aged 1;8 – 2;6, mean MLU 1.63). The participants were

recorded via audiotape interacting at home. English-speaking participants were observed while eating, playing with toys and reading books. Mandarin-speaking participants were observed playing with toys both indoors and outdoor, dressing, at mealtimes, reading books and engaging in social interactions. The speech was transcribed and coded to indicate first, whether the NPs received a generic or nongeneric interpretation and second, for the "semantic domain" ("animate", "artifact", "food" or "other") the NP belonged to.

Gelman & Tardif found that generics occurred in child-directed speech in both languages, but that generics were more frequent in English than in Mandarin and that generics were more frequent for animals than for all other domains of conversation.

In a second experiment, the authors examined caregiver speech in both English and Mandarin (Mandarin, n = 24 mean age of children 1;8.5; English, n = 24 mean age 1;8.20) in a more controlled environment. Participants were videotaped interacting while looking through a book, playing with basic toys and playing with mechanical toys. The English-speaking children were recorded in a laboratory setting. As in the first study, the participants' speech was transcribed and coded to indicate whether the NPs received a generic or nongeneric interpretation and for the "semantic domain" ("animate", "artifact" or "other")[34, 35] the NP belonged to.

As in the previous study, generic utterances were found to occur in child-directed speech in both languages and were more frequent in English than in Mandarin. Again, generic

[34] Because there are many cues to genericity in English, the English transcripts were further modified to remove pronouns, articles and morphological cues related to number (e.g. previously generic "you like buttons" became "like button" or previously nongeneric "what's the lion say?" became "what lion say?"). The rationale for modifying the English transcripts was that Gelman & Tardif wanted to determine how similar the Mandarin and English coding were and to ensure that the coding rubric did not over- or underestimate the number of generics in one of the languages. Upon recoding the modified English transcripts, more generics were identified than had been with the morphological cues present.

[35] "Food" was included in the "other" category in this study.

statements were more frequent in talk about animals than for all other domains, while nongenerics were more frequent for artifacts than for animals. There was no difference between English and Mandarin in frequency of nongenerics.

In a final study, the authors looked at adult interpretations of generics. Participants included 27 Mandarin-speaking adults and 23 English-speaking adults. They were presented with a questionnaire in their native language asking for the interpretation of 16 sentences from the English transcripts in Gelman & Tardif's Experiments 1 and 2[36]. Participants read actual utterances from the transcripts and were asked to judge whether the sentence referred to "one", "a few" or "most or any" members of the category. Gelman & Tardif found that adult speakers of both languages distinguished the utterances that had been generic from those that had been nongeneric, but English speakers made a more consistent distinction between the two than Mandarin speakers. On ten items, English speakers agreed 90%; Mandarin speakers agreed 90% on only three items. Gelman & Tardif explain this via the presence of articles, morphology (number and tense) in English.

The findings from this series of studies show that generic expressions are present in speech to children in two languages that differ in the surface forms generically interpreted NPs can take.

14) a. A kite goes flying in the sky.

b. Bunny rabbits don't make noises, do they?

c. Do you like buttons?

[36] The English sentences were translated into Mandarin for the Mandarin speakers.

The examples in (14), repeated here, are characterizing sentences produced by English speaking parents in this study. These examples show that parents are making use of the different NP forms that can be interpreted generically. Again, though, this series of studies does not explicitly report on the distributions of the different morphological determiner forms in child-directed speech.

1.6.4 Children's interpretations of generics

Pappas & Gelman (1998) and Gelman & Tardif (1998) give us some insight into the way caregivers use generic expressions in child-directed speech. These studies do not reveal much about the use of definite NPs, nor do they directly inform the question of whether or not children are getting the kind of input they need to be able to make use of the MH in learning about generic and referring language. However, they do show that children are hearing generic and referring expressions and, importantly, that parents are using both morphological forms of indefinites in the expression of genericity.

So while the previous studies do not allow us to look at the input directly, by looking at the kinds of interpretations children have for generically quantified indefinite NPs and comparing those interpretations to interpretations of NPs that are quantified with other quantifiers, we can gain some insight into whether children have been able to use the input to map the interpretations of morphological determiners onto the appropriate semantic determiners. If children have been able to make the mapping between semantic and morphological forms of determiners, then we should expect to see evidence that children can distinguish quantified and definite NPs from indefinites.

Gelman et al. (2002) examined children's use of quantifiers in inductive inferencing. They reasoned that if generic language allows speakers to make inferences about richly

structured categories, presumably children would be able to learn about categories through generic language. This study is relevant to the research presented in this dissertation because it examined children's interpretations of utterances with an indefinite subject (in this case a bare plural) that was used to make a statement about a property of a kind. Gelman et al. compared the interpretations of statements with bare plural subjects to statements with a universally- or existentially-quantified subject.

In this study, 37 child (mean age 4;7) and 36 adult participants were taught a novel property about an exemplar animal using one of three quantifiers: ∀ with overt *all*, ∃ with overt *some* or GEN with a bare plural NP without an overt quantifier. The rate of generalization to new exemplars of the category was measured. The authors also examined the typicality of the exemplars in generalization.

Participants were shown sets of six cards depicting a kind (e.g., fish). Three of the cards depicted highly typical members of the kind and three depicted less canonical exemplars of the kind. After participants saw a set of cards, the experimenter read a question about the set asking, *Which ones do x…?*. Half of the participants were provided with a clue (a novel property that was true of the kind) before seeing the animals and hearing the question; half were provided with the clue afterward. The clue was provided with one of the three quantifiers (i.e., *Fish/All fish/Some fish have small teeth in their throat*). Each participant heard three clues with each of the three quantifiers (blocked). Three of the clues used by Gelman et al. (2002) appear in example (60).

60) a. Bears/All bears/Some bears eat ants.

b. Birds/All birds/Some birds have a very sticky tongue.

c. Cats/All cats/Some cats have eyes that glow in the dark.

Importantly, both preschool children and adults distinguish GEN from both *all* and *some*, although age differences do exist in the interpretation of generic NPs. For children, GEN seems intermediate between *all* and *some* while for adults GEN is closer to *all*. These results show that children can appropriately interpret sentences with bare plural subjects as generic. Again, it should be noted that the properties taught to participants in this study are presented in characterizing sentences. The genericity of these utterances is dependent on the sentence itself[37] as is evident from the felicity of the generic interpretation in the following examples where the indefinite singular subject has been substituted for the bare plural that was actually used by Gelman et al. (2002).

61) a. A bear eats ants.

b. A bird has a very sticky tongue.

c. A cat has eyes that glow in the dark.

The previous study showed indirectly that children can distinguish generics from universally- and existentially-quantified statements by testing the degree to which a novel property taught to participants with GEN, *all* or *some* was generalized to the kind. Hollander et al.

[37] The clue sentences with *all* and *some* are not interpreted generically because the NPs are quantified by other operators (∀ and ∃).

(2002) was designed to test the meanings of GEN, *all* and *some* more directly and to compare 3- (n = 18, mean age 3;6) and 4-year-old children (n = 18; mean age 4;5) to adults (n = 38) using both a comprehension task and an elicited production task.

In the first of two studies, Hollander et al. (2002) manipulated both wording and property type. Participants received three blocks of questions, each using GEN, *all* or *some*, in one of three permutations of possible block orders. For 4-year-olds and adults, each block consisted of 12 questions: four about properties that were generically true of the kind (e.g., *Do animals/all animals/some animals eat?*), four about properties that were likely to be true of some members of the kind (e.g., *Do bears/all bears/some bears have white fur?*), and four about properties that were irrelevant for kind membership (e.g., *Are children/all children/some children made of feathers?*)[38].

The results from Hollander et al. (2002) that are relevant to the research here include the finding that properties that were generically true of the kind were most readily accepted, followed by properties that were true of only some members of the kind, followed by properties that were irrelevant to kind membership. Further, properties that were asked about using the existential *some* were more widely accepted than properties that were asked about with the bare plural subject, which in turn were more widely accepted than properties that were asked about with *all*. This indicates that participants were most discriminating in attributing properties to all category members. Three-year-old children were consistently sensitive to the property manipulation, but not to the wording manipulation. On the other hand, 4-year-olds were very nearly adult-like except that in response to questions with *some*, adults accepted more properties

[38] Hollander et al. called these "wide-scope", "narrow-scope" and "irrelevant" properties, respectively.

that were true of only some members of the category, whereas children did not distinguish these from generically true properties.

Hollander et al. (2002) also performed a production task in which 4-year-old children (*n* = 48, mean age 4;10) and 37 adults were asked to provide five properties[39] about 12 different kinds[40]. Each participant was assigned to the GEN, *all* or *some* condition and all of the properties they provided for the different kinds were to be true of GEN, *all* or *some* members of the category, respectively. Adults and 4-year-olds both distinguished GEN from *some*, but only adults distinguished GEN from *all*[41].

The last piece of previous research that I will discuss investigates children's use of NP form and pragmatic cues to interpret sentences as generic or nongeneric (Gelman & Raman 2003). In this series of five studies, the authors examined the role of NP form, comparing the use of bare plural (potentially generic subjects) to definite plural (nongeneric) subjects (e.g., *What color are dogs?* or *What color are the dogs?*). This study also investigated what the authors termed "pragmatic context". In other research, Gelman refers to "pragmatic context" as "contextual cues" discussed below. In this series of studies, Gelman & Raman investigated the congruence or mismatch of an anaphoric pronoun with the number depicted in each test item.

In their first two experiments, Gelman & Raman (2003) examined formal cues alone. Participants included 50 children (*n* = 16, mean age 4;7, *n* = 18, mean age 2;8, *n* = 16, mean age

[39] Children were asked to tell Zorg, a space alien, five things about each kind.
[40] The kinds consisted of natural kinds (e.g., birds), social categories (e.g., smart people) and artifacts (e.g. tables).
[41] An interesting, though not obviously relevant, finding was that both 4-year-olds and adults provided fewer physical properties and more action properties in the GEN condition than they did in either the *all* or *some* condition. Hollander et al. (2002) interpret this difference in property type in light of the claim in the literature that generics tend to be about essential properties. They suggest that such properties are less likely to be "…outward perceptual features and more likely to be actions and mental states" (p. 892).

3;6)[42] and 25 adults. Participants were shown a realistic drawing of two atypical exemplars of a kind (e.g., penguins, as atypical birds) and were asked a question about the dimension by which the exemplar was atypical (e.g., flying) in either a generic (e.g., *Do birds fly?*) or nongeneric (e.g., *Do the birds fly?*) form[43]. The responses that participants provided were classified as "specific," "category-wide" or "other". Gelman & Raman found that generic questions get generic responses, nongeneric questions get nongeneric responses. This effect was stronger among adults than the 4-year-olds, and relatively equal between 2- and 3-year-olds. They also found that category-wide responses were more common for animate items than inanimates for adults. This reflects the distribution of generics and nongenerics for animates and inanimates in previous research.

Gelman & Raman's second series of studies investigated the role of pragmatic cues[44] in generic interpretation. In these studies, two groups of 36 children (Study 2B: n = 12, mean age 2;8, n = 12, mean age 3;6, n = 12, mean age 4;7; Study 2C: n = 12, mean age 2;7, n = 12, mean age 3;4, n = 12, mean age 4;5) and 26 adults were presented with cards depicting one or two atypical members of a kind (e.g., tailless horse(s), tiny elephant(s), three-legged dog(s))[45] and asked questions about the atypical dimension. Each card was presented in one of three conditions: mismatch (one exemplar, plural NP: *Here is an elephant. Are they big or small?*), singular match (one exemplar, singular NP: *Here is an elephant. Is it big or small?*), plural match (two exemplars, plural NP: *Here are two elephants. Are they big or small?*). Participants received equal numbers of each of the test conditions. For adults, this consisted of all three test

[42] Experiment 1A examined adults and 4-year-olds; experiment 1B replicated 1A with 2- and 3-year-olds.
[43] Adults and 4-year-olds received 20 items; 2- and 3-year-olds received a subset of 12 items.
[44] Study 2A examined adults, studies 2B and 2C examined 2-, 3-, and 4-year-old children.
[45] Unlike the first pair of studies, the items had to be atypical in a way that is not typical of subtypes of the kind either. For instance, items could not be atypical in the way that penguins are atypical of birds because although most birds fly, it is typical of penguins that they don't.

conditions, while for children in experiment 2B, this consisted of the mismatch and singular match conditions and for children in experiment 2C, this consisted of the mismatch and plural match conditions.

As in the previous series of studies, responses were classified as "specific," "category-wide" or "other". As predicted, items that were presented in the mismatch condition received more generic responses than the two match conditions and items in both of the match conditions received significantly more nongeneric responses than items in the mismatch condition, except among the youngest group. For 2-year-olds, although there was a significant difference between generic and nongeneric responses between the mismatch and singular match conditions, there was no significant difference in generic and nongeneric responses between the mismatch and the plural match conditions. It is also noteworthy that there was a slight tendency for increases in generic utterances in plural-match condition among 4-year-olds. This is perhaps not surprising considering that in the plural match condition, it is possible that a speaker can make a claim about the kind, even in the presence of two exemplars of the kind (see discussion of example (54) in §1.4.1).

The results of this study show that children as young as two can distinguish between definites and indefinites as they are relevant to interpreting generic and referring language. Further, by age three, children are sensitive to certain pragmatic cues.

1.7 Conclusion

In this chapter, I have discussed the phenomena that have been described as generic and the problems determining what it means to be "generic". I then discussed the different components of grammar of genericity including quantification, interpretation of indefinites and Diesing's

(1992) Mapping Hypothesis. I then provided a learning algorithm and discussed how the previous literature on the acquisition of generics bears on that question. Finally, as a preface to the experimental chapters that address open issues in the literature, I provided evidence to show how treating genericity as a property of sentence grammar, better describes the linguistic signal. Such an analysis would also constrain the hypothesis space to one that children could use to learn to distinguish generic and referring language.

In the experimental chapters that follow, I test my claim that genericity is best understood in terms of the MH through experimentation that examines the three basic predictions the MH makes: first, the MH predicts that if the learner can find the indefinites, she will know how to interpret them; second, the MH predicts that children should have access to all the interpretations that are available for indefinites, regardless of what may be absent from the input; third, the MH predicts that if the grammar of genericity relies on syntactic structure for interpretation, then existentially interpreted indefinite subjects will prime other structures requiring VP-internal subjects (if syntactic priming exists as a phenomenon in child language[46]), while generically interpreted subjects will prime other structures requiring subjects interpreted outside of VP.

As I will show in Chapter 2, corpus data show that the learner can find the indefinites by examining the distribution of different morphological determiners and the syntactic position of NPs. Corpus data also show that indefinites outside of VP are interpreted as generic, but indefinites inside VP are interpreted as existential, consistent with MH. However, children do not get evidence in the input for the availability of existentially interpreted indefinite subjects so the findings of this study do not reveal what the status of the MH is in learning. It is possible that the

[46] See Bencini & Valian (2006), Huttenlocher et al. (2005), Savage et al. (2003, 2006) for studies that show syntactic priming in children for argument structure and passives, and Branigan et al. (2005), for NP structure.

MH leads to the correct understanding of indefinites, or, alternatively, it could be the case that the distribution in the input is the guiding force in learning. However, as I will show in Chapter 3, children know about existentially interpreted subjects despite a lack of evidence in the input. This supports the claim that the MH drives learning rather than being its output. Finally, in Chapter 4, I provide further experimental data in support of the MH by showing that existentially interpreted indefinite subjects prime other structures requiring VP-internal subjects, while generically interpreted subjects prime other structures requiring subjects interpreted outside of VP. Crucially, I show this by priming interpretations of target sentences that are similar to the primes only because they exhibit similar syntactic representations.

Chapter 2

2.0 Corpus data

As discussed in §1.6, previous work on the acquisition of generics (Gelman 2004; inter alia) has shown that parents use both types of indefinite NPs (bare plurals and indefinite singulars) in generic utterances, and has argued that the input by itself is uninformative regarding generic interpretation. However, the input is only uninformative if we expect children to be able to use the input to determine which NPs are generic. For this reason, I proposed in Chapter 1 that talking about genericity as a property of sentence grammar, rather than as a property of NPs, solves the learning problem faced by the previous literature.

Recall that the type of generic utterances children regularly hear are characterizing sentences with object-NPs, like the examples in (14), repeated here as (62). For generic utterances of this variety, being generic depends on an indefinite NP's location in the clausal architecture, rather than on the NP itself.

62) a. A kite goes flying in the sky.
b. Bunny rabbits don't make noises, do they?
c. Do you like buttons?

If we assume that children can take advantage of an innate piece of knowledge about the syntax-semantics mapping (the MH), they can solve the problem of learning about the difference between generic and referring language in the face of what would otherwise be a many-to-many

mapping between surface form and function. This is because the MH, discussed in detail in §1.3.3, predicts that if a learner can find the indefinites in his language, he will know how to interpret them: indefinite NPs that are inside VP are interpreted existentially, while indefinite NPs that are outside VP are interpreted generically (or within the scope of another overt quantifier or Q-adv). Once children have separated the definites from the indefinites, UG takes over and they can use the subject-direct object asymmetry predicted by the MH to gain insight into which NPs should be interpreted generically (all subjects of I-level predicates, and some subjects of S-level predicates) and which NPs should be interpreted existentially (some S-level subjects and all direct objects) on the basis of their structural position. This type of learning algorithm makes the complexity of learning about generics from the input more manageable.

Shifting the burden of genericity to the clausal architecture means that the learning problem no longer requires children to determine which NPs are generic; rather, it requires them to determine which NPs are indefinite and which are definite[47]. In order to do this, children would need to use the input to establish the mapping between morphological determiners in their language and semantic determiners provided by UG. Because the previous literature examined the acquisition of generics from a different perspective, the question of whether the input is sufficiently contentful that a child could make use of it to this end remains open.

In order to answer this question, two corpus studies were performed on child-directed speech, including a re-examination of the data obtained in Gelman & Tardif (1998), this time taking into account linguistic regularities in the input such as NP-form, distribution and information status, as well as predicate type. The features coded for in this database were chosen

[47] It also requires them to determine where NPs can sit in a Logical Form and when GEN can be inserted.

because these particular factors directly inform the learning algorithm proposed in §1.5 and are relevant to the appropriateness of the MH as a basis for my learning model.

In what follows, I motivate the feasibility of my learning algorithm by demonstrating that the input children hear is quite rich in the necessary ways. The input provides ample evidence for children to ascertain which morphological determiners line up with the indefinite and definite determiners in the semantics. The results of these corpus studies show strong support for the MH. Nearly without exception, subjects are interpreted outside of VP, as generics, and direct objects are interpreted within VP, as existentials. This generalization is true for both indefinite singular NPs and bare plural NPs. What is surprising is the absence of VP-internal subjects of S-level predicates. In other words, children are getting no evidence for the availability of interpretations like (38b), repeated here in (64), where the subject is interpreted inside VP. The input provides evidence for the availability of interpretations like those in (63) and in (64c), where indefinite subjects are interpreted outside VP.

63) Activists are passionate.

64) a. Activists are available

b. $\exists_x$ [*x* is an activist & *x* is available]

"There are activists available at some point in time."

c. $GEN_{x,s}$ [*x* is an activist & *s* is a time] [*x* is available at s]

"It is a (necessary) general property of activists that they are available."

2.1 Participants

2.1.1 Experiment 1

The data for this study were obtained from verbatim transcripts provided by Susan Gelman from her previous research (Gelman & Tardif 1998, inter alia), now available in the Tardif corpus in the CHILDES database archive (MacWinney 2000)[48]. The data in the current study consisted of transcripts from 22 monolingual English-speaking parent-child dyads engaged in a picture description task, which required the caregiver[49] and child to look through a picture book and talk about the pictures. Each page of the book depicted one exemplar of a particular kind and each kind was pictured only once in the book. The mothers and children were videotaped in a lab at the University of Michigan as they interacted in this and two other conditions – playing with simple toys and playing with mechanical toys. Each of the three sessions of play lasted 10 minutes. The contexts were in counterbalanced order. The parents had been instructed to play as they would at home and had been told that the researchers were studying how parents would play with their children with different kinds of toys. Results from Gelman & Tardif (1998) and Gelman & Raman (2003) showed that there were significantly higher rates of generic usage in the picture description condition than in either of the other two conditions. For this reason only the transcripts from the picture description task were examined here. The present study investigated child-directed speech (i.e., the parents' utterances) only. The children in these transcripts ranged in age from 1;5 – 1;11, mean age, 1;9.

[48] Because these transcripts are available in the Tardif corpus in CHILDES, they will subsequently be referred to as the Tardif transcripts.
[49] In these transcripts, all the caregivers were mothers.

2.1.2 Experiment 2

The data for this study were obtained from the Brown corpus (Brown 1973), now available through the CHILDES database archive (MacWinney 2000). Adam was one of three participants in a longitudinal study that was completed between 1962 and 1966 and reported on in Brown 1973. As a follow-up to the transcripts from Gelman & Tardif 1998, Adam's first two transcripts were entered into the database. Adam was slightly older than the children in the corpus used in Experiment 1: 2;3.4 in the first transcript and 2;3.18 in the second. The data were collected in Adam's home and reflect normal interactions between Adam and (in the two transcripts used in Experiment 2) his mother.

2.2 Coding procedure

Coding procedures were the same for both corpora, with the exception of the recording of information status, which will be discussed in §2.2.2, below. Each common noun (i.e., pronouns were excluded) that was produced by a caregiver was identified. Information about each noun and the utterance it occurred in were entered into a Microsoft Access database designed by the author. Each noun's database entry included the sentence the noun occurred in and the verb it it was an argument of. Additionally, an NP's entry included whether it was singular or plural, definite or indefinite, and its position in the syntax. Each entry also included whether the sentence the NP occurred in was characterizing or particularizing as well as the interpretation the NP received in the sentence. Finally the type of predicate (S- or I-level) was also recorded. A sample of the database form used to record information is included in Figure 4.

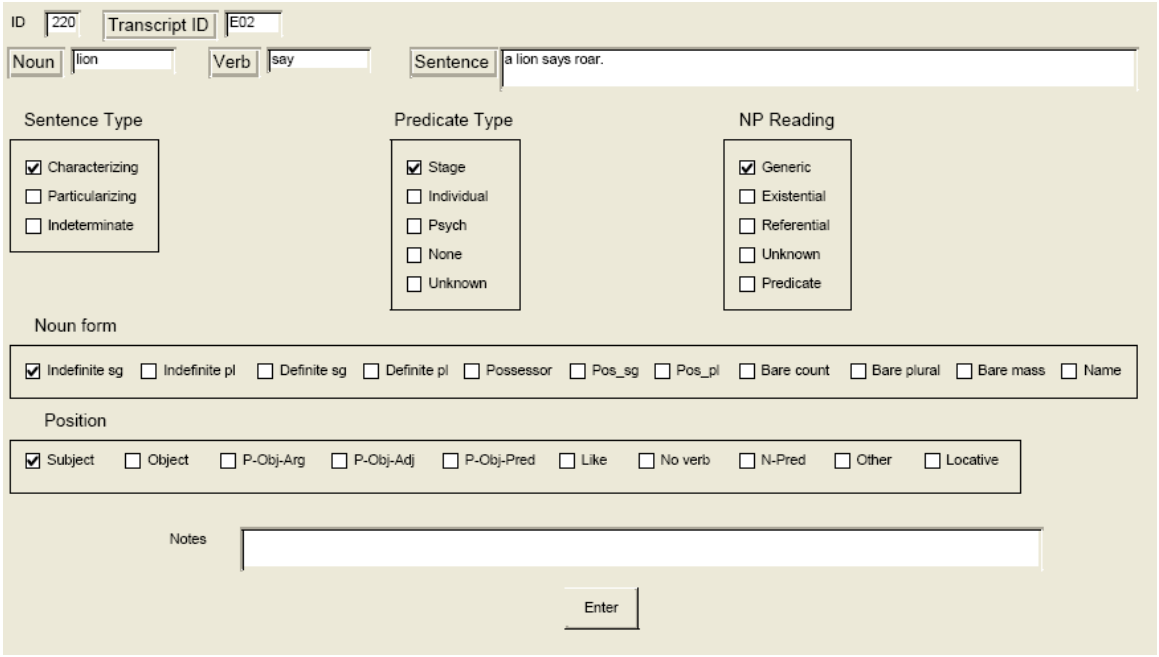


Figure 4. Sample database form

The author was the only coder, so every effort was made to be consistent and principled in labeling each item. Classifications were based on taxonomies found in the literature, primarily those discussed in *The Generic Book* (Krifka et al., 1995) and in *The Essential Child* (Gelman, 2003). The discussion of the features coded for in this database is divided into three sections based on what aspect of the MH or learning algorithm the particular feature is relevant to: formal subcategories, information status and semantic subcategories. The features in the different subcategories and how they are used by the learning algorithm are explained in each of the following subsections.

2.2.1 Formal subcategories

The type of information reflected in the formal subcategories, in connection with information status (see §2.2.2, below), enables the learner to distinguish indefinites from definites. This represents the first stage of the learning algorithm.

2.2.1.1 NP-form

The MH applies only to the interpretation of indefinite NPs. Therefore, a child's first task in the proposed learning algorithm is to separate the definite NPs from the indefinite NPs. In order to do this, a child needs to be able to map the morphological determiners that are available in her language onto the semantic determiners that she is provided innately. This can be done by attending to the morphological form of the determiner that is used with the NP when the noun is uttered. Recall from the discussion in §1.5 that the intuition behind this piece of the algorithm lies in the different functions of definite and indefinite NPs. Being an "indefinite" means that something introduces a discourse referent so the learner can make the inference from seeing an NP with *a*/*an* as the determiner being used to introduce a discourse referent to the knowledge that NPs with *a*/*an* as the determiner have Kamp-Heim semantics and represent the indefinites.

For an adult speaker, the mappings are already explicit. Adults know, for example, that *a* maps onto "indefinite" and *the* maps onto "definite". In labeling NP-form, I am using my knowledge as an adult speaker of English to classify the forms. The process a learner would go through is somewhat different. A learner would hear an NP used for the first time in a discourse to introduce a discourse referent, and would have to make a note of the morphological determiner that was used with it. After hearing enough NPs with *a* or *an* used to introduce discourse referents, the child would know that these morphological determiners have "indefinite"

semantics. Likewise, for definites, the learner would make the inference from hearing an NP with *the* as the determiner being used to refer back to a referent that is already part of the discourse to the knowledge that NPs with *the* as the determiner have definite semantics.

By recording the NP form in the database, we can examine the distribution of NP forms used on the first and subsequent mentions of particular NPs. We should see that the morphological forms used to introduce discourse referents are *a*/*an*/Ø. However, since only definites can be used anaphorically, they are expected to be used more as second mention than are bare plurals and indefinites. A definite NP may also be used for the first mention of a discourse referent if the context allows the speaker to point to a particular, unique individual. Thus, we expect definites to occur relatively equally often as first and second mention, but we expect an asymmetry for indefinites. We may also find that there are significantly fewer tokens of definites than indefinites if, as Fisher & Tokura (1995) found, caregivers tend to use fewer definite NPs in speech to children than in speech to other adults.

2.2.1.2 Syntactic position

The syntactic distribution of the NPs is also relevant to record in the database because the MH predicts different interpretations for subjects (of some S- and all I-level predicates) and direct objects by virtue of the fact that subjects may be interpreted in IP and direct objects are interpreted in VP[50]. This information is also relevant for the learning algorithm because like NP-form, syntactic position can help the learner distinguish the definites from the indefinites. As noted in §1.5, canonical subjects are more likely to represent information that has already been mentioned in the discourse, and as a consequence, are more likely to be definite than are non-

[50] As noted in §1.3.3, the objects of psych predicates are an exception to this.

subjects. Indefinites canonically introduce new information (which is more natural following given information), and are more common later in the sentence (Prince 1992).

2.2.2 Information status

In my discussion of the two previous subcategories, I have linked their utility to their interaction with information status. It is by using NP-form and syntactic position in combination with information status that a learner can determine which NPs are indefinites. Thus, this information was recorded too.

While it was not recorded in the database's entry form in Figure 4, after all the nouns for each transcript had been entered into the database, the number of times each NP had been uttered in that transcript was recorded by hand for each transcript in the Tardif transcripts. The first time an NP was mentioned, it was labeled "1", the second time, "2", the third time, "3", and so on. By this classification, an NP is discourse new on its first mention and discourse old on its subsequent mentions. Because each kind was pictured on only one page in the picture book, it is possible to say that each time a particular NP was used it referred to the same discourse referent.

The number of utterances intervening between two mentions was also recorded. This was done in the event that an NP could be considered "new" again after a certain interval. However, because the parent-child dyads went through the book different numbers of times and spent different amounts of time on the different pictures in the book (depending on the child's interest), it was impossible to devise an appropriate algorithm for an NP's subsequent occurrence as "new" and the entire session was considered a single discourse.

Adam's transcripts were not coded for mention. There were two main reasons for this. First, Adam's transcripts were much longer than the 10-minute per participant in the Tardif

transcripts. The first of the Adam transcripts used here represented an hour-long recording; the second, a two-hour recording. Ten-minute intervals were not noted in the transcript, so it was unclear how to divide the transcripts into analogous discourse units in order to be able to compare mention in these transcripts to mention in the Tardif transcripts. Second, for reasons of time and resources, the sample is much smaller. Excluding all but the first and second mentions of NPs would have excluded too many NPs from the database. The comparison is admittedly far from perfect, but there are still interesting comparisons to make.

2.2.3 Semantic subcategories

Once the learner has been able to distinguish indefinite and definite NPs, he can use the type of information reflected in the semantic subcategories to determine which NPs get a generic interpretation.

2.2.3.1 Predicate-type

As noted in §1.4.1, the predicate in a generic utterance usually occurs in the non-progressive present tense. Tense thus indicates perhaps the most obvious contribution of the predicate; however, the type of predicate an NP occurs with is also important because the MH makes different predictions about the structural positions that are available to the indefinite subjects of different kinds of predicates. For instance, the subjects of S-level predicates can be interpreted in IP or VP and the sentences can receive a generic or existential interpretation, respectively. The indefinite subjects of I-level predicates can be interpreted only in IP and such sentences can only receive a generic interpretation. The indefinite subjects *and* direct objects of psychological predicates are interpreted outside of VP and these sentences receive a generic reading.

It is relevant to record this type of information because it can give the learner insight into what interpretations are available for particular subjects, constraining the hypothesis space[51].

2.2.3.2 Sentence type

Recall from the discussion in Chapter 1 (§1.2, 1.4.1, 1.5) that particularizing sentences are tied to a particular occurrence or point in time. It is common for the tense of the predicate that appears in such sentences to denote an event in the past (e.g., "Did you eat an apple today?", transcript E01) or, alternatively, an event that is currently taking place with the use of the progressive (e.g., "What's this little girl doing?", transcript E03). Characterizing utterances, on the other hand, are notably lacking any links to a particular time or place. They report general properties of individuals or situations (e.g., "A lion says roar", transcript E01). Utterances that are used to label or categorize are also characterizing (e.g., "Uh, that's a soccer ball", transcript E01). Utterances classified as indeterminate consisted of, primarily though not exclusively, utterances without a verb (e.g., "Apple", transcript E01).

This information was recorded in the database because it reflects the presence (or absence) of spatiotemporal information that could prevent (or allow) the generic operator from being introduced into the representation. Particularizing sentences do not allow the generic operator to be introduced into the representation because they contain information that makes the sentences true or false at particular times or locations (e.g., *last week*, *at school*). As noted in §1.2, particularizing sentences can only be interpreted generically when they contain a kind-NP.

[51] It should be noted that there were utterances in the corpus that consist of what seems to be labeling, where the caregiver simply says the noun (e.g. "Apple.", transcript E01). The predicate type of such utterances is classified as NONE.

However, the reader is reminded that such generic utterances do not occur in the sample of child-directed speech under discussion here.

Based on the examples above, it is clear that not all characterizing sentences are generic; however, all of the generic utterances that occur in this sample are characterizing sentences. Characterizing sentences don't contain information that links the utterance to a particular time or place, so they allow the generic operator to be inserted into the representation, whether or not it actually is. The difference between characterizing sentences that are generic and those that aren't is that the generic operator has been introduced into the representation of the former and not the latter. For generic utterances of this variety, being generic depends on an indefinite NP's location in the clausal architecture, rather than on the NP itself.

2.2.3.3 NP-Reading

The analysis of genericity that has been put forward in this dissertation treats genericity as a property of the sentence grammar. Only kind-NPs can be generic on their own; object-NPs, which are under discussion here, are interpreted generically only if they are located in a particular position in the syntax. It is important to continue to make the distinction between genericity via an NP and genericity via the whole sentence. I have recorded the interpretation of each NP in the database in order to look at how the other features recorded in the database are linked to generic utterances, but I maintain the point of view that genericity, with the exception of kind-NPs, is the result of an NP's position in the clausal architecture. It is still necessary to record an NP's interpretation, though, because ultimately the learner has to decide on an interpretation for the NP, even if she uses the syntax to do it.

An NP's categorization was determined by a combination of cues including morphosyntactic, semantic and pragmatic information. In this corpus, an NP could only be classified as generic if it had the appropriate form and occurred in the appropriate sentence type. For the purposes of this corpus study, the NP forms that could be generic were bare plurals, indefinite singulars and bare mass nouns[52]. The only sentence type that was considered generic was characterizing, though, as noted above not all characterizing sentences are generic. This category serves to distinguish generic from nongeneric characterizing sentences.

NPs that are labeled "Existential" are those whose existence is necessary, but which are not specific (e.g., "Did you eat *an apple* today?", transcript E01), while those that are specific are "Referential" (e.g., "No, don't pull *my finger* out!", transcript E01). NPs that are used as predicates were labeled "Predicate" (e.g., "Uh, that's *a soccer ball*", transcript E01). Also in the corpus are utterances that consist of what seems to be labeling or categorizing where the caregiver simply says the noun, without a verb (e.g., "*Apple*", transcript E01). NPs that occurred in such contexts were also labeled "Predicate".

2.3 Results

2.3.1 Experiment 1

The total number of nouns in the database was 2672. Seven hundred and seventy-six NPs were excluded from the analyses because they occurred without a verbal predicate, and therefore without much syntactic information. This left 1896 nouns for analysis. Over 65% of these nouns, 1236 of them, represent a first (804) or second (432) mention. Because the parent-child dyads

[52] Although the definite singular can be interpreted generically, it was excluded from that classification here for the reasons noted above in §1.2.

often went through the book more than once and the amount of time spent on the different pictures in the book was dependent on each child's interest, only first and second mentions were analyzed. Thus the analyses below are based on 1236 NPs, unless otherwise noted.

What is quite noticeable from a first look at the data is that, as shown in Table 1, there are far fewer generics than there are existentials and referentials. In fact, there are nearly five times as many existentials (n = 291) and over seven times as many referentials (n = 426) as there are generics (n = 60). What is particularly interesting about this difference, though, is the distribution of the interpretations across different syntactic positions. First of all, referentials are more evenly distributed across subject and object position than are either the generics or the existentials. Of the referential NPs, 15% occur in subject position and 22% occur in direct object position. With both generics and existentials, there is a marked subject-direct object asymmetry; however, the asymmetry is in different directions. Nearly all of the generic NPs are subjects (93%), whereas there is only one existential subject and there are lots of existential direct objects (57%). A 3x2 χ^2 test of independence with NP-reading (generic vs. existential vs. referential) and position (subject vs. direct object) as factors shows that the observed distribution is different from the expected distribution (χ^2 = 188.76, $p \leq 0.0001$). All the cells vary from what is expected, however the largest deviations from expected are in the generic and existential cells. In particular, the cells with the largest deviation from what is expected are the Generic x Subject cell, which has 193% more observations than expected, z = +8.44 and the Existential x Subject cell, which has 100% fewer observations than expected, z = -7.25. The Referential cells vary less from what is expected (Referential x Subject, 31.3% more observations, z = +2.21; Referential x Object, 14.6% fewer observations, z = -1.51).

Reading (%)		Subject (%)	Object (%)	Other (%)
Generic	60 (5)	56 (93)	4[53] (7)	0 (0)
Existential	291 (24)	1 (0)	165 (57)	126 (43)
Referential	426 (34)	65 (15)	92 (22)	268 (63)
Other	459 (37)	15 (3)	50 (11)	394 (86)

Table 1. NP interpretation by grammatical function

The findings reported above suggest that the input is not distributed randomly. Let us now go through the data and determine whether the input is informative for the learning algorithm proposed in §1.5.

2.3.1.1 Learning Algorithm: Phase 1

Recall that the first stage of the learning algorithm requires the child to separate definite NPs from indefinite NPs. This can be accomplished by attending to the form used when the noun is mentioned. This is because the two NP forms have different functions. While either an indefinite or a definite NP may be used for the first mention of a discourse referent, only definite NPs may be used anaphorically and so are expected to be used relatively more often the second time an NP is mentioned than are indefinite NPs. However, as noted in §2.1.1.1, definite NPs may also be used the first time a discourse referent is mentioned if (roughly) conditions of uniqueness are met. Thus, we expect definites to occur relatively equally often as first and

[53] These four tokens of generic objects reflect four direct objects of psych verbs, which, unlike most direct objects, are interpreted generically.

second mention, but we expect an asymmetry for indefinites. This is, in fact, exactly what we see in the input. Table 2, below, shows a difference in distribution for these two classes of NP.

NP form (%)			First Mention (%)	Second Mention (%)
Indefinite	664	(54)	455 (69)	209 (31)
Definite	280	(23)	158 (56)	122 (44)
Other	292	(24)	191 (65)	101 (35)

Table 2. NP-form by mention

Definites occur in relatively equal percentages across first (56%) and second (44%) mention, while indefinites display a very different pattern. Among indefinites, they are twice as likely to occur as the form a noun takes on its first mention (69%) than they are as the form it takes on its second mention (31%). A 2x2 χ^2 test of independence with NP-form (indefinite vs. definite) and mention (first vs. second) as factors shows that the observed distribution is different from the expected distribution ($\chi^2 = 12.13$, $p = 0.0005$). The cell with the largest deviation from what is expected is the Definite x Second Mention cell, which has 23.8% more observations than expected, $z = +2.35$. This finding is in line with our expectations that definites will be used more often as a second mention than will indefinites.

An examination of the raw numbers in Table 2 might seem to provide evidence against the predicted distribution of NPs because, numerically, there are more instances of second-mention indefinite NPs than there are instances of second-mention definite NPs. However, this is

only because there are more than twice as many indefinites as there are definites in this sample[54]. Table 3 illustrates the weighted distribution of NP-forms across first- and second-mention. If we assume that there were equal numbers of indefinite and definite NPs in the sample, there would be 122 second-mention definites and only 88 second-mention indefinites.

NP form (%)	First Mention (%)	Second Mention (%)
Indefinite 280	192 (69)	88 (31)
Definite 280	158 (56)	122 (44)

Table 3. Weighted NP-form by mention

A 2x2 χ^2 test of independence on the weighted values with NP-form (indefinite vs. definite) and mention (first vs. second) as factors shows that the observed distribution is different from the expected distribution (χ^2 = 8.3, p = 0.004). The cell with the largest deviation from what is expected is the Definite x Second mention cell, which has 15.7% more observations than expected, z = +1.61. This finding, too, supports the prediction that definites will be used more often as a second mention than will indefinites.

Children can further distinguish the definites from the indefinites by looking at where in the sentence the NP occurs. Recall Prince (1992)'s findings about subjects and objects and their relationship to information status: canonical subjects are more likely to represent information that has already been mentioned in the discourse making them more likely to be definite than

[54] It is noteworthy that children are getting more than twice as many indefinite NPs in their input than definite NPs. This echoes the findings from Fisher and Tokura (1995) showing that mothers were significantly less likely to use the definite with children than they were with adults.

non-subjects. Indefinites introduce new information which is more natural following given information, making them more common later in the sentence. Again, this prediction is borne out in the data.

NP form (%)	Subject (%)	Object (%)	Other (%)
Indefinite 664 (54)	60 (9)	160 (24)	444 (67)
Definite 280 (23)	48 (17)	53 (19)	179 (64)
Other 292 (24)	30 (10)	98 (34)	164 (56)

Table 4. NP-form by grammatical function

In Table 4, we can see that definites are more evenly distributed between subject and direct object position with 17% of definites occurring in subject position and 19% of definites occurring in direct object position. In contrast to this, there are nearly three times as many indefinites in direct object position (24%) as there are in subject position (9%). A 2x2 χ^2 test of independence with NP-form (indefinite vs. definite) and position (subject vs. direct object) as factors shows that the observed distribution is different from the expected distribution (χ^2 = 11.83, p = 0.0006). The cell with the largest deviation from what is expected is the Definite x Subject cell, which has 39.8% more observations than expected, z = +2.32. This finding is in line with our expectations that definites will be used more often early in the sentence will indefinites.

2.3.1.2 Learning Algorithm: Phase 2

Once the child has separated definites from indefinites, she knows which nouns are relevant to the MH. The definite NPs can be ignored and the distribution of indefinite singular and bare plural NPs can be examined more carefully. As discussed in Chapter 1 and in §2.1 of this chapter, the difference between the sentences in (63) and (64), repeated below, can be explained by the LF position of the subject. The sentence in (63) has only one potential subject position: [Spec, IP]. Material contained in IP, but not VP, is mapped into the restrictive clause of the covert generic operator; thus, the sentence in (63) is interpreted generically. The sentence in (64a) has two available subject positions: [Spec, IP] and [Spec,VP]. If the subject is interpreted inside VP, in the lower subject position, the sentence will be interpreted existentially (as in (64b)). However, if the subject is interpreted outside of VP, the sentence will instead be interpreted generically (as in (64c)).

63) Activists are passionate.

64) a. Activists are available.

b. $\exists_x$ [*x* is an activist & *x* is available]

"There are activists available at some point in time."

c. $GEN_{x,s}$ [*x* is an activist & *s* is a time] [*x* is available at *s*]

"It is a (necessary) general property of activists that they are available."

As shown in Table 5, there is a striking difference in where generics and existentials occur in the input. With a single exception, subjects are interpreted outside of VP, as generics,

and direct objects are interpreted within VP, as existentials. This generalization is true for both indefinite singular NPs and bare plurals. What is surprising is the near absence of existential subjects, which are predicted to be possible with S-level predicates. In other words, children are not getting evidence for the availability of interpretations like (64b), where the subject is interpreted VP-internally.

To examine this in greater detail, let us first consider the distribution of indefinite singular NPs illustrated in the top half of Table 5. A 3x2 χ^2 test of independence with NP-form+Position (indefinite subject vs. indefinite direct object vs. indefinite other) and interpretation (generic vs. existential) as factors shows that the observed distribution is different from the expected distribution ($\chi^2 = 215.63$, $p < 0.0001$). The cell with the greatest deviation from what is expected is the Subject x Generic cell, which has 416.3% more observations than expected, $z = +11.87$. Because the conditions for a χ^2 test of independence are not met by the bare plural data, a Fisher exact test of probability was run, where the treatments were NP-form+Position (bare plural subject vs. bare plural direct object vs. bare plural other) and the outcomes were interpretation (generic vs, existential). The results show that the outcomes are not independent of the treatments ($p < 0.0001$).

Reading / Position		Generic (%)	Existential (%)
Indefinite			
Subject	45 (8)	42 (93)	1 (2)
Object	120 (21)	0 (0)	92 (77)
Other	396 (71)	0 (0)	87 (22)
Bare plural			
Subject	14 (14)	14 (100)	0 (0)
Object	40 (39)	3 (8)	37 (93)
Other	49 (48)	0 (0)	20 (41)

Table 5. Interpretations by NP-Form and position[55]

2.3.2 Experiment 2

Perhaps the most notable feature of the data in Table 6 is that there are far fewer generics than there were in the Tardif transcripts (see Table 1). This is to be expected based on previous findings that generics are used more often in talk about natural kinds, which represent a great deal of what parents and children were looking at in the Tardif transcripts examined in Experiment 1. However, a similarity between the transcripts in Experiment 1 and these here is that there are still far more existentials and referentials than there are generics. Furthermore, we see similar distributions in these data of the different NP readings across syntactic positions. First of all, referentials are more evenly distributed across subject and direct object position than are

[55] The three examples of generic objects are with psychological predicates, a finding that is expected based on the MH. Recall that in Table 1, there were four generic objects (all occurring with psychological predicates). The additional token of a generic object, not included in Table 4, which looks only at indefinites, bare plurals and definintes, was a bare mass noun (*spaghetti*).

the existentials[56]. Of the referential NPs, 29% occur in subject position and 32% occur in direct object position. There is, again, a marked subject-direct object asymmetry among the existentials. There are only three existential subjects and lots of existential direct objects (75%).

Because the conditions for a χ^2 test of independence are not met by these data, a Fisher exact test of probability was run, where the treatments were subject and direct object and the outcomes were interpretation (generic, existential, referential). The results show that the outcomes are not independent of the treatments ($p < 0.0001$).

Reading (%)		Subject (%)	Object (%)	Other (%)
Generic	2 (0)	1 (50)	1[57] (50)	0 (0)
Existential	79 (11)	3 (0)	59 (75)	17 (25)
Referential	306 (44)	89 (29)	99 (32)	118 (39)
Other	311 (45)	23 (7)	26 (8)	262 (85)

Table 6. NP interpretation by grammatical function for Adam

First and second mention were not examined for the reasons explained in §2.2.2; however, the distribution of the different NP forms across syntactic position can still be examined. The data in Table 7 show that definites occur in equal percentages in subject (23%) and direct object (28%) position, while there are nearly 14 times as many indefinites in direct object position (34%) as there are in subject position (2%). A 2x2 χ^2 test of independence with NP-form (indefinite vs. definite) and position (subject vs. direct object) as factors shows that the

[56] It is hard to make the same claim about generics, given that there are only two.
[57] This token of a generic object reflects one direct object of a psych verb.

observed distribution is different from the expected distribution (χ^2 = 18.66, $p < 0.0001$). The cell with the largest deviation from what is expected is the Indefinite x Subject cell, which has 75.4% fewer observations than expected, z = -2.84. The Definite x Subject cell has 45% more observations than expected, z = +2.2. These findings are in line with the findings from Experiment 1, and confirm the prediction that definites will be used more often early in the sentence than will indefinites.

NP form (%)			Subject		Object		Other	
Indefinite	127	(19)	3	(2)	43	(34)	81	(64)
Definite	150	(21)	35	(23)	42	(28)	73	(49)
Other	421	(61)	66	(16)	100	(24)	255	(60)

Table 7. NP-Form by grammatical function for Adam

Although the data on NP-form and mention cannot be taken into account for this set of transcripts, the findings regarding the distribution of definite and indefinite NPs reflect the results of the first corpus study. This suggests that in everyday parent-child interactions, a child would be able to determine which NPs are definites and which are indefinites. Recall from the proposed learning algorithm that once the child knows which NPs are indefinite, he can observe their syntactic distribution to determine the interpretation of the sentences in which they occur.

Table 8 illustrates the distribution of interpretations by NP-form and syntactic position for the NPs in the Adam corpus. The conditions for a χ^2 test of independence are not met by these data so a Fisher exact test of probability was run, where the treatments were NP-

form+Position (indefinite subject vs. indefinite direct object vs. indefinite other) and the outcomes were interpretation (generic vs, existential). The results show that the outcomes are not independent of the treatments ($p = 0.010$).

Reading / Position			Generic		Existential	
Indefinite						
Subject	6	(5)	1	(17)	0	(0)
Object	34	(27)	0	(0)	34	(100)
Other	86	(68)	0	(0)	61	(71)
Bare plural						
Subject	1	(3)	1	(100)	0	(0)
Object	14	(44)	1	(7)	11	(79)
Other	17	(53)	0	(0)	0	(0)

Table 8. Interpretations by NP-Form and position for Adam

The sample of bare plural NPs is even smaller than the sample of indefinite singular NPs for this corpus. Again, the conditions for a χ^2 test of independence are not met by these data so a Fisher exact test of probability was run, where the treatments were NP-form+Position (bare plural subject vs. bare plural direct object vs. bare plural other) and the outcomes were interpretation (generic vs, existential). The results show that the outcomes are independent of the treatments ($p = 0.154$). It is possible that with a larger sample, and more bare plurals, a significant difference would emerge.

2.4 Discussion

These findings suggest that the input is informative in determining an NP's interpretation. The results of these corpus studies show that there are statistical and distributional cues in parental input that can aid the child in the acquisition of generic and referring language. The input to children is structured in such a way that a child could make use of it as described in the learning algorithm. Based on the available input, the child should be able to use mention and syntactic position to separate the definites from the indefinites. From there, she can again use the distribution in the syntax to determine which indefinites are generic and which are existential.

For both the corpora examined here, almost all subjects are interpreted outside of VP, as generics, and direct objects are interpreted within VP, as existentials. This generalization is true for both indefinite singulars and bare plurals. In other words, children are getting only very limited evidence for the availability of interpretations like (64b), where the subject is interpreted inside VP. This is surprising because the MH predicts ambiguity for indefinite subjects of S-level predicates. This finding raises the question of whether the input constrains the interpretations available to children, or if it merely serves to confirm that the generic interpretation is available. If the MH is part of children's grammars, children should have access to interpretations that are not in their input. This question is investigated in Experiment 3 in Chapter 3.

Chapter 3

3.0 Interpretation of bare plural subjects

The findings from Experiments 1 and 2 suggest that the input is informative in determining how an NP is interpreted in an utterance based on its syntactic position at LF. These corpus studies show that the learner can find the indefinites by examining the distribution of different morphological determiners and the syntactic position of NPs. Corpus data also show that indefinites outside of VP are interpreted as generic and indefinites inside VP are interpreted as existential, consistent with the MH. But as noted in §1.6, the results of these studies do not reveal anything about how children interpret indefinite NPs in utterances that are intended to be generic versus those in utterances that are intended to be referential. Further, children do not get evidence in the input for the availability of existentially interpreted indefinite subjects. The findings of Experiments 1 and 2 do not reveal what the status of the MH is in learning. It is possible that the distribution in the input constrains learning, such that children are not able to interpret indefinite subjects existentially because they do not occur in the input, or that the MH leads to the correct understanding of indefinites because children know that *both* interpretations are available despite the limited evidence available in the input. The goal of the experiment reported in this chapter was to test these two hypotheses.

In order to test these two alternatives, this experiment capitalized on the fact that indefinite subjects of S-level predicates are ambiguous between a generic and an existential interpretation. Use of this ambiguity permits us to determine which interpretation is available to children (if both are not) and which is preferred (if both are). Using the Truth Value Judgment

Task (TVJT) methodology and bare plural subjects[58], this experiment presented participants with utterances like the following, which are ambiguous out of context.

65) a. Crocodiles are in the desert area.

b. $\exists_x$ [*x* is a crocodile & *x* is in the desert area]

c. GEN_x [*x* is an crocodile] [*x* is in the desert area]

As discussed in §1.3.3, the indefinite subject in (65) is ambiguous between a generic and an existential interpretation, depending on where it is interpreted. The sentence in (65b) says that there are crocodiles that are in the desert area, but the sentence in (65c) indicates that it is a general property of crocodiles that they are in the desert area (i.e., maybe they live there).

If participants are presented with a context that makes the test sentence in (65a) true on the semantic interpretation in (65b) and they judge the test sentence to be true, this tells us that the participants have access to the existential interpretation in (65b). For instance, if participants are presented with a situation in which some crocodiles at a zoo have gone over to the desert area to eat lunch and they accept the test sentence (65a), this reveals that these participants have access to the existential interpretation. If participants are presented with a context that makes the test sentence in (65a) true on the semantic interpretation in (65c) and they judge the test sentence to be true, this tells us that the participants have access to the generic interpretation in (65c). For instance, if participants are presented with a situation in which crocodiles live in the desert area at the zoo (regardless of where the crocodiles are at the telling of the story) and they accept the test sentence (65a), this reveals that they have access to the generic interpretation. It is important

[58] Bare plural NPs are less restricted in their use as generics than indefinite singulars (Cohen 2001b).

to note that an affirmative response does not indicate that participants do not have access to the other logically possible interpretation, only that they have access to at least the interpretation that they agreed with. If, however, participants reject a test sentence, it reveals that they do not have access to the interpretation that would have made the test sentence true.

Adults controls are expected to have access to both the logical interpretations available for sentence (65a). They would be expected to accept the test sentence in (65a) at rates of acceptance near ceiling both in a context that would make the sentence true on its existential interpretation, as in (65b), and in a context that would make it true on its generic interpretation, as in (65c). The predictions for children are more complex. Based on the findings from Experiments 1 and 2, if children are limited to using the input to learn about the meanings of indefinite subjects, children would be expected to interpret bare plural subjects generically and judge test utterances like (65a) as false when the context makes the sentence true on its existential interpretation. Children would be predicted to accept test sentences only when the context makes them true on their generic interpretation. If children reject test sentences that are true on their existential interpretation, this would suggest that the input strongly influences children's interpretation of NPs and that while it does give them insight into genericity, the full range of NP meanings must be learned through additional means. Such a finding would provide evidence against the MH's being part of children's grammars at this stage of development.

Alternatively, if the MH is part of children's grammars, children should have access to both interpretations of indefinite subjects, despite what is in their input. If this is the case, children should behave like adults and accept the test sentence in (65a) both in contexts that make it true on its existential interpretation and in contexts that make it true on its generic interpretation.

As a control measure, unambiguously existential and generic utterances were presented to participants via *there*-sentences and sentences with an overt *usually*, respectively. Both adults and children were expected to accept these control sentences at high rates in the appropriate contexts and to reject them in contexts where they are false. Additionally, filler items were included to insure participant attentiveness.

3.1 Participants

This study examined two populations of native English speakers: normal adults and normally developing 4-year-old children.

3.1.1 Children

The children (n = 72) in this study were recruited from two sources. The first source consisted of children recruited from a database of families who had expressed interest in participating in studies on children's cognitive development at Northwestern University's Project on Child Development. These children participated in this study on site, on the Northwestern campus. For their participation, each child was given the choice of a small toy or book. Children in the second group were recruited at preschools and child care centers in the northern suburbs of Chicago, including Evanston and Northbrook. These children participated in the study in a quiet room of their preschool during the school day. To thank the school for allowing us to run our studies at the school, a per-child donation was made to the school for academic supplies or activities.

The number of children included in the analyses was 52 (31 girls, 21 boys) ranging in age from 4;0 to 4;9; mean age 4;6. An additional 20 children were run in the study but were excluded from the final analysis for one or more of the following reasons: age (n = 1), bilingual (n = 1),

inability/unwillingness to complete the experiment (n = 3), response bias or innattention during the experiment[59] (n = 15).

3.1.2. Adults

The adults (n = 52) were recruited from the Northwestern University Department of Linguistics subject pool. This population consisted of undergraduate students at Northwestern who were enrolled in 200-level linguistics classes, which require students to participate in linguistics experiments for course credit. Prior to their participation in the study, adult subjects were asked to fill out a language background questionnaire. Only monolingual native speakers of English with no reported history of relevant hearing or language problems were included in the analyses. An additional eight subjects were excluded for the following reasons: language background (n = 4), inattention to task (n = 1), experimenter error (n = 3).

3.2 Method

This experiment used the version of TVJT developed by Stephen Crain and colleagues (Crain & McKee 1985, Crain & Thornton 1998). This methodology generally requires two experimenters: one to act out and describe a series of stories that the participants and a puppet will watch, and one to control the puppet, to which the participants will ultimately respond. To begin, participants are introduced to the puppet and told that the puppet is just learning and needs their help. Any number of reasons can be provided to establish the puppet's need for help and the participant's expertise (i.e., the puppet is a space alien who is just learning English or the puppet is just learning, period, and sometimes gets confused). This part of the methodology is

[59] As measured by incorrect responses to filler items.

particularly important for child subjects as they are often very eager to correct the puppet's mistake (so they can help him learn) where they might be reluctant to tell an adult experimenter that she is wrong. Furthermore, Crain and colleagues' innovation was to make the task one which relies on a system of reward and punishment rather than requiring the "yes" or "no" response that earlier versions of the task used. The idea behind using reward or punishment is that there is usually some humorous element involved (e.g., the child can reward a puppet's "correct" response with a cookie and punish an "incorrect" response by making the puppet eat something gross or do push-ups). This serves to engage the child in the task and it makes both "right" and "wrong" fun so that children are less likely to be biased against saying "no"[60].

During the task, participants watch a series of stories and hear the puppet's follow-up statement about each story. Participants judge whether or not the puppet has felicitously spoken and they mete out reward or punishment, accordingly.

The TVJT is a particularly useful method for evaluating a child's comprehension of complex syntactic constructions because children are not required to make any metalinguistic judgments about the constructions under investigation. Instead, children are required only to say whether a puppet's test sentence corresponds to what happened in a story they just observed. This is, presumably, relatively easy for children to accomplish and allows the experimenter to examine a child's grammar without the limitation of performance factors (e.g., memory). For these reasons, the TVJT is ideally suited as an experimental method to examine the syntactic representations that children have at an abstract level of grammatical structure.

[60] Children's yes-bias has long been observed in the literature, but see Fritzley and Lee (2003) for a recent experimental investigation of the phenomenon.

3.3 Stimuli

3.3.1 Test stories

The test stories in this study took advantage of the interpretive ambiguity of bare plural subjects (discussed in §1.3.3) by, first, presenting what was usually true in the world of the story and then, second, demonstrating to participants that a state of affairs different from the usual state of affairs was true on this occasion.

In each of the test stories, there were three different kinds of entities that were thematically related to the story in analogous ways (e.g., three different kinds of animals at the zoo, three different acts at the circus). These three kinds were located in three different locations within the context of the story (e.g., monkeys are in the jungle, crocodiles are in the river area, zebras are in the savanna). The main character in the story always did something with each kind of entity, whether it was simply to look at the members of the kind or to interact with them[61]. Throughout the telling of the story, the relevant entities that were to subsequently appear in the puppet's test sentences were always referred to with definite NPs (e.g., the/these/my crocodiles). All participants heard the same series of test and filler stories, counterbalanced for order.

3.3.2 Test sentences

The between subjects factor was sentence form. There were three different sentence forms used in this experiment: *bare plural* (*BP*), *there*, and *usually*. The first of these sentence form types was comprised of sentences with a bare plural subject. These were the sentences that tested participants' access to the two different interpretations available for indefinite subjects: existential or generic. The other two sentence forms (a *there*-sentence and a sentence with an

[61] Within each story, the main character's interaction with each kind was the same.

overt expression of genericity, *usually*) acted as control sentences to measure participants' access to these two interpretations in unambiguous sentences. The *there*-sentences should be true on their existential interpretation and the *usually*-sentences should true on their generic interpretation. Each participant heard test sentences of only one of these forms.

The within subjects factor was context. Each sentence form had two versions so that within the context of a story, each ambiguous *BP*-sentence was true on its existential interpretation or on its generic interpretation (but not both) and only one version of the unambiguous controls was true. In other words, a particular *BP*-sentence in the context of the story may be true on its existential reading and false on its generic reading *or* it may be true on its generic reading and false on its existential reading. These two interpretations of bare plural subjects represent the contexts. In the EXT context, an utterance is true on its existential interpretation. In the GENT context, an utterance is true on its generic interpretation. The unambiguous controls were true in only one context. The *there*-sentence was true in the EXT context and false in the GENT context, while the *usually*-sentence was true in the GENT context and false in the EXT context. Each participant heard a block of three sentences that were true on their generic interpretation and a block of three sentences that were true on their existential interpretation.

The six permutations of the test sentences are illustrated in the examples below.

66) *BP*

a. Crocodiles are in the desert area.

b. Crocodiles are in the river area.

67) *There*

a. There are crocodiles in the desert area.

b. There are crocodiles in the river area.

68) *Usually*

a. Crocodiles are usually in the desert area.

b. Crocodiles are usually in the river area.

The dependent measure was %-acceptance. This was chosen as the dependent measure because if a participant accepts a test sentence in the EXT context, it does not mean that she does not have access to the generic interpretation of the sentence as well. The only claim that can be made from a positive response to the puppet's test sentence is that that particular reading is available to the participant. It is only if a test sentence is rejected that we know a participant only had access to one of the logically possible interpretations. Additionally, using %-acceptance as the measure makes the interpretation of the results more intuitive.

3.3.3 Test items

To demonstrate how the stories and the test sentences come together, it will be helpful to go through one of the stories as an example. I will go through the "Alice at the zoo" story step by step.

As this story about a zoo begins, participants can see four landscapes (three with animals in them, one that is empty), a zookeeper and Alice (from *Alice in Wonderland*), who is visiting the zoo.

Figure 5. Alice at the zoo, story begins

The first part of each story involves establishing what is usually (generically) true within the world of the story. This aspect of each of the stories is intended to avoid the problem of different world knowledge backgrounds between subjects. In this story, the zookeeper character describes to Alice where all the different animals live: the monkeys live in the jungle, the crocodiles live in the river area and the zebras live in the savanna. The zookeeper goes off to work while Alice decides which animals she wants to see first. She decides on the monkeys. While Alice is over by the jungle area looking at the monkeys, the zookeeper notices that the river where the crocodiles live is filthy and he needs to clean it.

At some point in each story, an event occurs that makes what is generically true false on this particular occasion. In other words, the actual state of affairs at the end of the story differs from the usual state of affairs. In the case of the zoo story, the zookeeper sets the occurrence of the requisite exceptional event in motion when he notices that the river area needs to be cleaned. At this point, the zookeeper makes a decision to clean the river right away and says that he'll need to move the crocodiles to do it. He moves the two crocodiles from the river habitat they normally live in into the empty desert area. He then goes back to the river to clean.

Alice decides to go see the zebras in the savanna next and after watching them for a while decides she will go see the crocodiles in the river area. To her surprise, when she arrives at the river area, the crocodiles are gone and the zookeeper is cleaning. He tells her that the river was dirty and that he had to move the crocodiles in order to clean it. She walks away, dejected, but perks up as she notices that that there's something in the desert area. She sees that it's the crocodiles and is happy to see them even if she does not get to see them swim.

Figure 6. Alice at the zoo, story ends

Consider again the examples in (66) – (68). The sentences with bare plural subjects in example (66) are ambiguous. In principle, (66a) could either mean that there are crocodiles in the desert area right now *or* that crocodiles are usually found in the desert area. The same is true of (66b), modulo location. The sentences in (67) describe where the crocodiles are right now, whether that be the desert or the river, and the sentences in (68) describe where the crocodiles usually are, again, whether that be the desert or the river.

66) a. Crocodiles are in the desert area.

$\exists_x$ [x is a crocodile & x is in the desert area]

"There are crocodiles in the desert area."

$GEN_{x,t}$ [x is a crocodile & t is a time] [x is in the desert area at t]

"It is a general property of crocodiles that they are in the desert area."

b. Crocodiles are in the river area.

$\exists_x$ [x is a crocodile & x is in the river area]

"There are crocodiles in the river area."

$GEN_{x,t}$ [x is a crocodile & t is a time] [x is in the river area at t]

"It is a general property of crocodiles that they are in the river area."

67) a. There are crocodiles in the desert area.

$\exists_x$ [x is a crocodile & x is in the desert area]

"There are crocodiles in the desert area."

b. There are crocodiles in the river area.

$\exists_x$ [x is a crocodile & x is in the river area]

"There are crocodiles in the river area."

68) a. Crocodiles are usually in the desert area.

$USUALLY_{x,t}$ [x is a crocodile & t is a time] [x is in the desert area at t]

"Usually for crocodiles they are in the desert area."

b. Crocodiles are usually in the river area.

$USUALLY_{x,t}$ [x is a crocodile & t is a time] [x is in the desert area at t]

"Usually for crocodiles they are in the river area."

However, once we consider the context of the story, only one of the interpretations available for the bare plural sentences is compatible with the story and only one of the two sentences in each of (67) and (68) is true. In the context of what happens in the story, example (66a) is true only in the EXT context and (66b) is true only in the GENT context. Example (67a) is true because at the end of the story there are crocodiles in the desert area, but (67b) is false because only the zookeeper is in the river area. Example (68b) is true because crocodiles live in the river area, but (68a) is false because crocodiles are not usually found in the desert area.

66) a. Crocodiles are in the desert area.

$\exists_x$ x is a crocodile & x is in the desert area

"There are crocodiles in the desert area."

~~$GEN_{x,t}$ [x is a crocodile & t is a time] x is in the desert area at t~~

~~"It is a general property of crocodiles that they are in the desert area."~~

b. Crocodiles are in the river area.

~~$\exists_x$ x is a crocodile & x is in the river area~~

~~"There are crocodiles in the river area."~~

$GEN_{x,t}$ [x is a crocodile & t is a time] x is in the river area at t

"It is a general property of crocodiles that they are in the river area."

67) a. There are crocodiles in the desert area.

$\exists_x$ x is a crocodile & x is in the desert area

"There are crocodiles in the desert area."

b. There are crocodiles in the river area.

~~$\exists_x$ x is a crocodile & x is in the river area~~

~~"There are crocodiles in the river area."~~

68) a. Crocodiles are usually in the desert area.

~~$\text{USUALLY}_{x,t}$ [x is a crocodile & t is a time] x is in the desert area at t~~

~~"Usually for crocodiles they are in the desert area."~~

b. Crocodiles are usually in the river area.

$\text{USUALLY}_{x,t}$ [x is a crocodile & t is a time] x is in the desert area at t

"Usually for crocodiles they are in the river area."

The stories and test sentences together fulfill two important requirements of the TVJT: plausible dissent and falsification. The pattern of truth conditions illustrated for the test sentences satisfies TVJT's condition of falsifiability because within the context of the story, each ambiguous test sentence is true on only one interpretation and only one member of each pair of control test sentences is true. Crucially, the crocodiles have been observed in both locations during the story. This invites participants to consider both potential locations in the story and thus both the existential and generic interpretations for the test sentences. This satisfies the requirement of plausible dissent because if the story had been different in an obvious way that is still consistent with the story, it would have been appropriate to reject the claim. A complete list of the test items including story summaries and test sentences is included in Appendix A.

3.3.4 Experimental conditions

In a TVJT, the truth conditions of the test sentence can be manipulated in one of two ways. Either the stories can differ across subjects, or the test sentences themselves can vary. In this study, there was only one version of each story and all the participants observed the same six test stories[62]. What varied between participants in this case were the test sentences.

Participants were tested with one of three sentence forms (bare plural, *n* = 28 children, 28 adults; *there*-sentence, *n* = 12 children, 12 adults; *usually*-sentence, *n* = 12 children, 12 adults) in one of two orders (Order EX-FIRST: sentences that were true in the EXT context followed by sentences that were true in the GENT context, *n* = 26 children, *n* = 26 adults or order GEN-FIRST: sentences that were true in GENT context followed by sentences that were true in EXT, *n* = 26 children, *n* = 26 adults). The most natural way to discuss the difference between the two orders is to do so from the perspective of the test condition (*bare plural*), rather than from the perspective of the two control conditions (*there*- and *usually*-sentences). I will discuss the ordering for the *BP*-sentences first and subsequently explain how this bears on the control sentences.

Recall that the goal of this study was to determine whether it is the distribution in the input that drives the acquisition of generic and referring language, such that children are not able to interpret indefinite subjects existentially, or, whether it is the case that children have access to both interpretations of indefinites because the MH is part of their grammars. In a TVJT, for a given test sentence, if a participant has access to the interpretation that is true in context, they should accept the test sentence as true; however, if they only have access to the interpretation that is false in context, they should reject the puppet's test sentence as false. Thus, by presenting participants with ambiguous test sentences that are true on only one of their interpretations in the

[62] Participants also observed the same filler stories.

context of the story, it is possible to determine whether they have access to a particular interpretation. If participants are presented with a test sentence that is true on its existential interpretation and they accept it, it shows that the existential interpretation is part of their grammars; however, it they reject it, we know that that representation is not allowed by their grammar.

Each of the two versions of the *BP*-sentences is ambiguous, but the context of the story makes only one interpretation of each version true, as described in §3.3.3. In order EX-FIRST, participants heard a block of three test sentences that were true in EXT, followed by a block of three test sentences that were true in GENT. The participants in order GEN-FIRST heard their test sentences in the opposite order: a block of three test sentences that were true in GENT, followed by a block of three test sentences true in EXT. The stories were blocked to protect against potential interference effects of the two available interpretations. If there were discrepancies in the rates of acceptance across ordering conditions, the first block of each order would reveal the participants' unbiased and unprimed rates of acceptance for that interpretation.

The task of dividing the stories into EXT and GENT blocks was accomplished by using the location of the relevant kind after the exceptional event for the EXT context (as in (66a)) and the usual location in the story of the relevant kind of entity for the GENT context (as in (66b)). In other words, (66a) is an EXT context test sentence because at the end of the story the crocodiles are in the desert area. The interpretation of the sentence that is consistent with the context is the existential interpretation, below.

66) a. Crocodiles are in the desert area.

$\exists_x$ x is a crocodile & x is in the desert area

"There are crocodiles in the desert area."

Example (66b) is a GENT context test sentence because the crocodiles are normally found in the river area, even though on this particular occasion they are not (because the zookeeper had to clean the river that they live in). The interpretation of this sentence that is consistent with the context is the generic interpretation, below.

66) b. Crocodiles are in the river area.

$GEN_{x,t}$ [x is a crocodile & t is a time] x is in the river area at t

"It is a general property of crocodiles that they are in the river area."

The two versions of the control sentences (*there*- and *usually*-sentences) are not ambiguous, but they have different truth conditions and only one version is true within the context of the story. The *there*-sentences are true only in the EXT context (as in (67a)) and the *usually*-sentences are true only in the GENT context (as in (68b)). Like the *BP*-sentence condition, the two blocks appear in both orders for these conditions, and the task of dividing the stories into EXT and GENT blocks was accomplished in the same way.

Within each order, the relative order of the six individual test stories was counterbalanced and each of the two blocks contained two filler stories.

3.4 Procedure

The procedure in this experiment varies slightly between children and adult controls. Both groups are introduced to the puppet (in this case a Jerry Garcia doll), but the procedures diverge at this stage.

3.4.1 Children

Children were tested individually in a quiet room. Some parents chose to sit in the room to observe the experiment, but they were instructed to remain quiet and not to say anything about the stories or the child's responses while the experiment was in progress. With the exception of one parent, who answered her cell phone during an experiment, parents complied with this request, occasionally telling their child to watch the story or encouraging the child to answer Jerry if the child turned to look at them.

After children were introduced to Jerry, they were told that Jerry was just learning and needed their help because he sometimes gets confused. They were told that they would watch some stories with Jerry and that at the end of each story, Jerry would say something that he thought was true of the story. If Jerry was right, the participant got to reward him by giving him a plastic cookie and if Jerry was wrong, the participant got to "punish" him by making him do sit-ups (which Jerry didn't mind too much because he wanted to get in shape). Participants then heard a simple warm-up story to familiarize themselves with the task. Once the child demonstrated understanding of the reward/punishment procedure, the child and Jerry then watched a series of stories (both test items and fillers) that were acted out and described by one of the experimenters. At the end of each story, Jerry provided a brief recap of the story and a test sentence. Participants evaluated the truth or falsity of the test sentence he provided.

Children were often asked follow-up questions as a way of determining whether their affirmative/negative responses to Jerry reflected the aspect of the grammar being examined in the study. If children said Jerry was wrong, they were always asked to tell Jerry why he was wrong and to tell him what the real answer was so Jerry could learn. Although it seems somewhat more plausible to ask a participant to correct the puppet than to reiterate why Jerry was right, children were also asked to tell Jerry what happened in stories that he had described correctly in order to ensure that Jerry "didn't forget what happened".

3.4.2 Adults

Adults were tested in a quiet room seated at a table, individually or in groups of up to three people. Since they were sometimes tested in groups and an oral response could bias other participants' interpretations of the test sentence on each item, or subsequent items, all adults wrote their judgments of Jerry's test sentences on an answer sheet.

After adult participants were introduced to Jerry, they were told that the experiment was designed to be told to children. This accounted for the use of props like Cookie Monster and Elmo and the fact that they would be evaluating what a puppet said. Because the ruse of a puppet who is just learning was not necessary for adults, the puppet need not be as animated or interactive and one experimenter was sufficient.

Participants were told that they would watch a series of stories and that at the end of each story, Jerry would say something that he thought happened. The participants were instructed that if they thought Jerry was right, they should circle "true" on their answer sheet; if they thought he was wrong, they should circle "false". Adults were told that they always had to justify their

responses, whether they thought what Jerry said was true or false, and they were given space to do so on the answer sheet.

The experimenter told the stories with Jerry sitting off to the side. When it was time for Jerry to provide his test sentence, he was moved to the center of the table directly in front of the experimenter. The experimenter held Jerry as Jerry provided a brief recap of the story and the test sentence was spoken.

3.5 Results

The results will be divided into two sections. The first section will consist of the 12 adults and 12 children in each of the *there*-sentence and *usually*-sentence conditions and the first 12 adults and children from the *BP*-sentence condition. The second section will concern only the results from the *BP*-sentence condition.

3.5.1 BP-sentences vs. control sentences

The analyses in this section compare the *BP*-sentences to the two forms of control sentences. The *BP*-sentences were the sentences that tested participants' access to the two different interpretations available for indefinite subjects and the control sentences (*there*-sentences and sentences with an overt expression of genericity, *usually*) measured participants' access to each interpretation in unambiguous sentences. The *there*-sentences should be true on their existential interpretation and the *usually*-sentences should true on their generic interpretation. Each participant heard test sentences in only one of these forms.

The child data reported here are from 36 children (20 girls, 16 boys) ranging in age from 4;0 – 4;9 years old (mean age 4;6). The adult data reported here are from 36 Northwestern undergraduates.

Correlated samples t-tests show that, as expected, for both children and adults, the *usually*-sentences were accepted significantly more often when they were presented in GENT contexts than when they were presented in EXT contexts (children, $t(11) = -10.34$, $p < 0.001$, one-tailed; adults, $t(11) = -812409.61$, $p < 0.001$). For children, the *there*-sentences had a non-significant tendency to be accepted more often when they were presented in EXT contexts than when they were presented in GENT contexts ($t(11) = 1.65$, $p = 0.063$, one-tailed). For adults, the effect was significant ($t(11) = 8.65$, $p < 0.001$, one-tailed). *BP*-sentences were not accepted at significantly different rates in the two contexts for children ($t(11) = -0.56$, $p = 0.587$, two-tailed), but for adults, there was a non-significant tendency to accept more *BP*-sentences in the EXT context than in the GENT context ($t(11) = 1.86$, $p = 0.089$, two-tailed).

The results for all three sentence forms are displayed in Figure 7 for children and Figure 8 for adults.

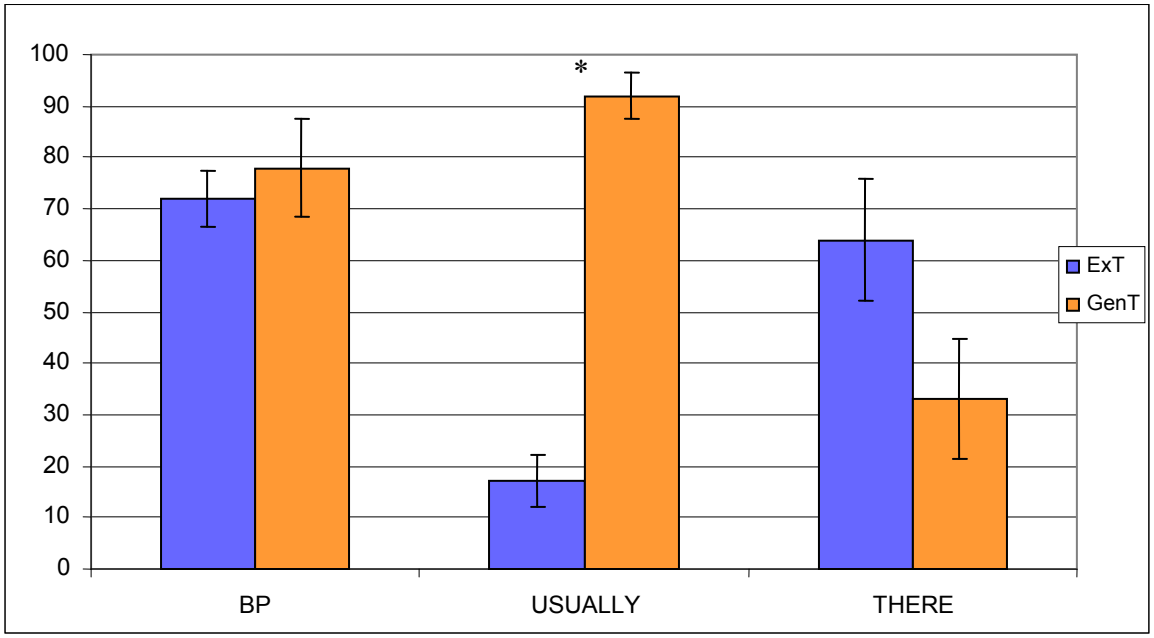


Figure 7. Children's acceptance of context by sentence form ** p < 0.001*

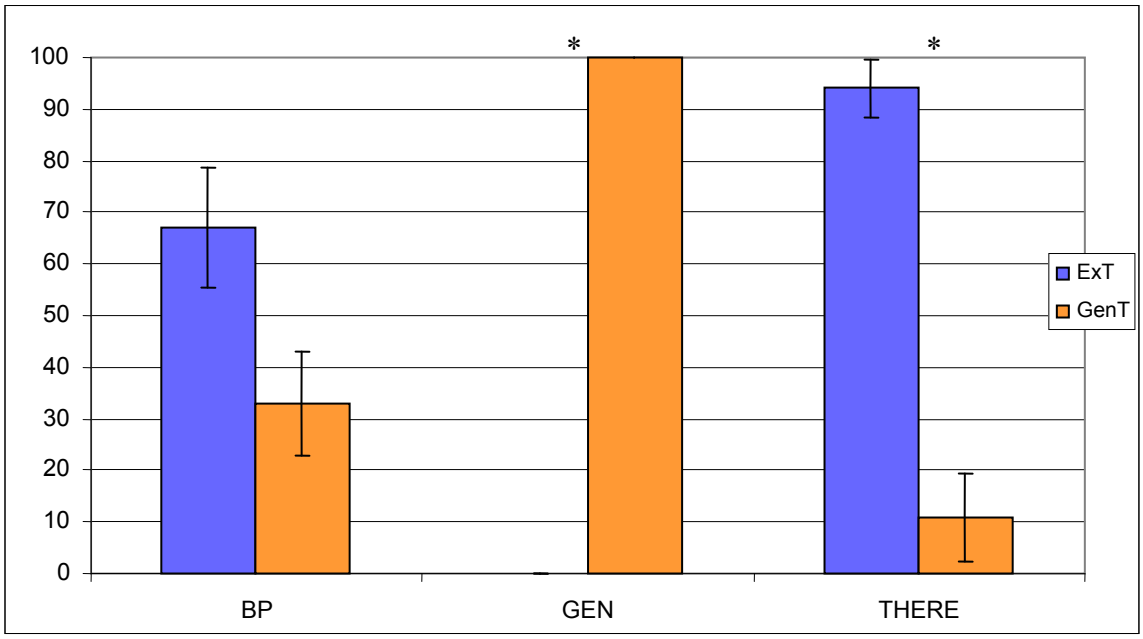


Figure 8. Adult's acceptance of context by sentence form[63] ** p < 0.001*

[63] Adults were at ceiling (100%) uniformly accepting *usually*-sentences in GENT contexts and rejecting them in EXT contexts.

These results show that both children and adults have access to generic and existential interpretations in unambiguous contexts. In other words, when participants are presented with an overt, unambiguous expression of genericity like *usually*, they interpret it generically significantly more often than they interpret it existentially. Likewise, when they are presented with an overt, unambiguous expression of existence like a *there*-sentence, they interpret it existentially more often than they interpret it generically. The difference is significant for adults, and approaches significance for children.

The interpretation of *BP*-sentences in the two contexts will be discussed in detail in §3.5.2, below.

3.5.2 BP-sentences

As described above in §3.5.1, the two control conditions were indeed the controls that they were expected to be. However, because relatively few data points were available for each subject in each of the conditions, the test condition with ambiguous bare plural subjects showed more variability and a more fine-grained analysis of the interpretation these sentences required more data. Additional subjects were run in this condition only, bringing the totals to 28 usable children and 28 usable adults.

The child data reported here are from 28 children (17 girls, 11 boys) ranging in age from 4;0 – 4;10 years old (mean age 4;6). An additional 16 children were run in the *BP*-sentence condition but were excluded for the following reasons: age (n = 1), inability/unwillingness to complete task (n = 3), bilingual (n = 1), missing two or more filler items (n = 11). The adult data reported here are from 28 Northwestern undergraduates. An additional five adults were run in the *BP*-sentence condition but were excluded as non-native speakers of English (n = 2), or due to

experimenter error (n = 3). For both children and adults, the dependent measure was the proportion of trials in which the participant judged the puppet's statement to be true.

Recall that adults were predicted to accept both interpretations of bare plural subjects at high rates, since they should have access to both logically possible interpretations of bare plural subjects. If children rely on the input to determine the interpretations of bare plural subjects, they were predicted to accept only sentences that are true on the generic interpretation of bare plural subjects, rejecting the existential interpretation because it was absent from their input. On the other hand, if the MH is active in children's grammars, they should accept sentences that are true on their existential interpretation as well as sentences that are true on their generic interpretation.

An Analysis of Variance (ANOVA) with three factors with two levels each: age (adult vs. child) x order (EX-FIRST vs. GEN-FIRST) x context (EXT vs. GENT) was conducted. The overall results are depicted graphically in Figure 9. There is a main effect of age ($F(1, 26) = 6.50$, $p < 0.017$), with children accepting significantly more test sentences than adults, and a main effect of context ($F(1, 26) = 17.9$, $p < 0.0003$), with test sentences in the EXT context being accepted significantly more often than sentences in the GENT context. There is an Age x Context interaction ($F(1, 26) = 10.6$, $p < 0.003$), with children and adults accepting the two contexts at different rates. Finally, there is a three-way Age x Order x Context interaction ($F(1, 26) = 11.3$, $p < 0.002$) showing that children and adults differ significantly in the rates of acceptance of the two contexts in the two orders. This result is driven by the fact that children always accept the test sentence type that they hear first at a higher rate than they accept the test sentence type they hear second whereas adults always accept the test sentences in EXT at a higher rate than the test sentences in GENT.

Before I discuss these findings in more detail, it is important to note two aspects of the overall picture. First, children *do* have access to the existential interpretation of bare plural subjects, despite the absence of this interpretation in their input. Second, adults do not show particularly high rates of acceptance for the test sentences in either context (EXT or GENT), and in fact strongly disprefer the generic interpretation, accepting test sentences in GENT only 40% of the time in EX-FIRST and only 24% of the time in GEN-FIRST.

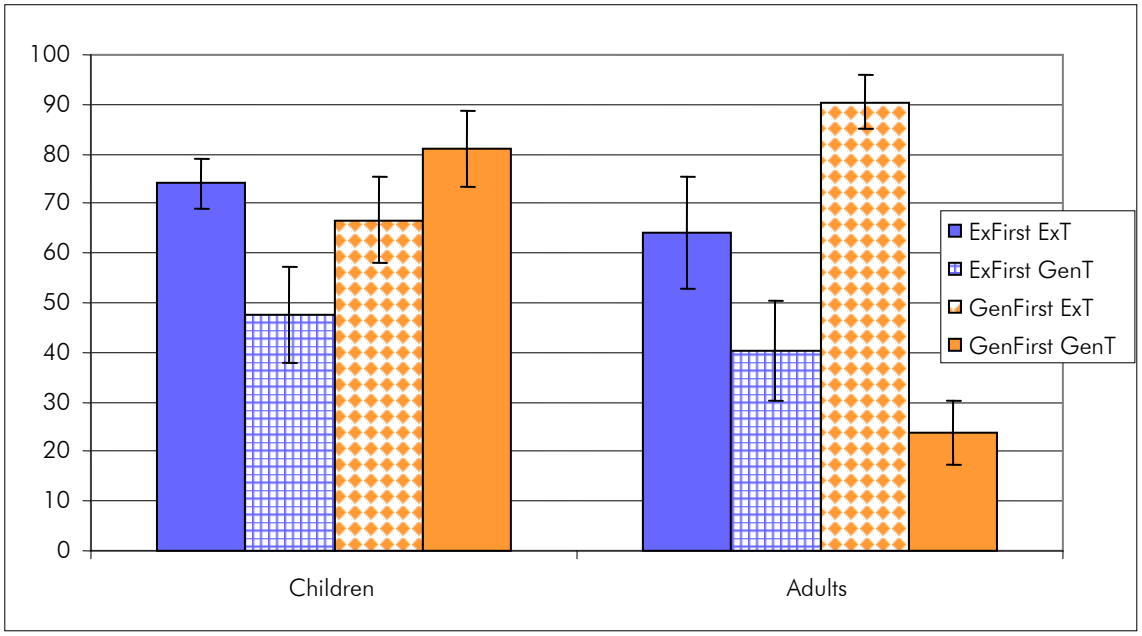


Figure 9. Mean rates of acceptance by age, order and context.

Collapsing across orders, an independent samples t-test comparing rates of acceptance for adults and children shows that children did not differ significantly from adults in their acceptance of sentences in EXT, accepting them 70% of the time compared to 77% for adults ($t(54) = -0.85$, $p = 0.399$, two-tailed); however, children accepted test sentences 64% of the time in GENT, while adults accepted them only 32%. The difference in rates of acceptance in GENT easily reached significance ($t(54) = +3.54$, $p < 0.008$, two-tailed). This is illustrated in Figure 10 below.

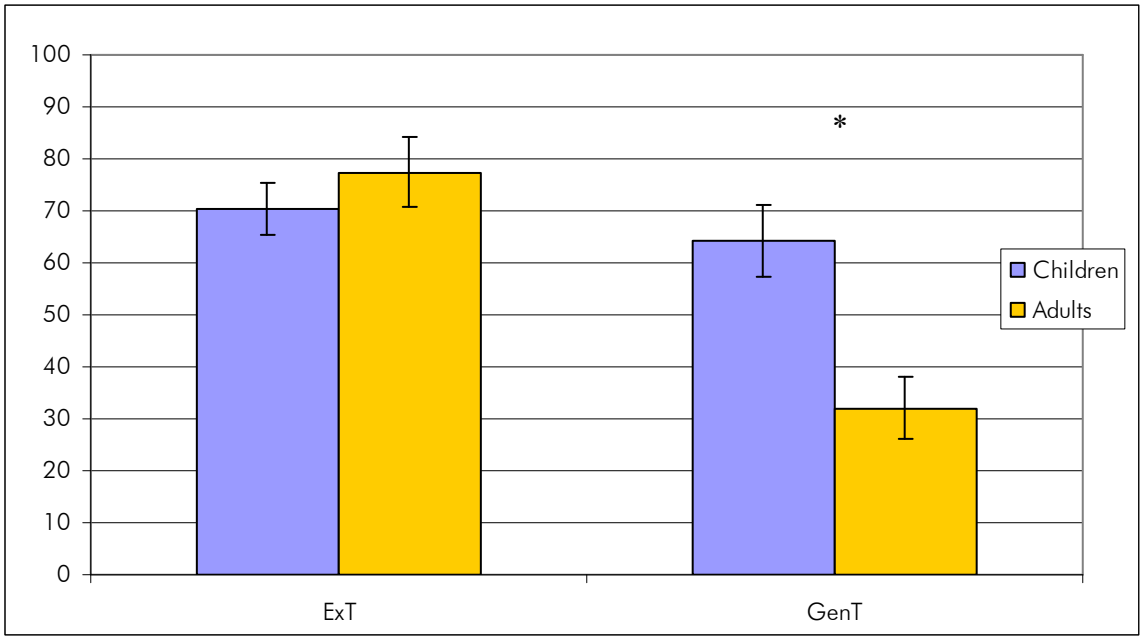


Figure 10. Mean rates of acceptance of test sentences by context and age ** p < 0.01*

Because the findings for children and adults were so different, and because the results were unexpected, it will be useful to go through the data in greater detail for each group separately.

3.5.2.1 Children

Considering the results for the children alone, as shown in Figure 10, children accepted 70% of the test sentences in ExT and 64% of the test sentences in GenT. A paired samples t-test shows that there is no significant difference between these two rates of acceptance ($t(27) = 0.67$, $p = 0.508$, two-tailed). Single-samples *t*-tests comparing the rate of acceptance of each context to chance shows that the children in Ex-first are significantly different from chance for the ExT context, but not for the GenT context (Ex-first, block 1: $t(13) = 4.66$, $p < 0.001$; block 2: $t(13) = -0.235$, $p = 0.818$; two-tailed). Children in Gen-first are significantly different from chance for the GenT context and marginally different from chance for the ExT context (Gen-first, block 1:

$t(13) = 4.10$, $p = 0.001$; block 2: $t(13) = 1.98$, $p = 0.07$; two-tailed). However, recall that each participant heard a block of three test sentences that were true on one of the two logically possible interpretations (existential or generic), followed by three test sentences that were true on the other interpretation. Participants in EX-FIRST heard three EXT sentences in the first block and three GENT sentences in the second block. The orders were reversed for participants in GEN-FIRST. Participants in GEN-FIRST heard three GENT sentences in the first block and three EXT sentences in the second block.

Children in EX-FIRST accepted 74% of the test sentences in the EXT context and 48% of the test sentences in the GENT context and children in GEN-FIRST accepted 67% of the sentences in EXT and 81% of the sentences in GENT. A 2x2 repeated measures ANOVA with order (EX-FIRST vs. GEN-FIRST) and context (EXT vs. GENT) as factors shows that while there was no main effect of order ($F(1, 26) = 2.88$, $p = 0.102$) or context ($F(1, 26) = 0.53$, $p = 0.473$), there was an interaction between the two. As displayed in Figure 11, children's rates of acceptance of test sentences in GENT are significantly lower in EX-FIRST than they are in GEN-FIRST ($F(1, 26) = 6.16$, $p < 0.020$). A one-way ANOVA for independent samples comparing the mean rate of acceptance of utterances in GENT when the GENT context was in block 1 and when it was in block 2 also shows significance ($F(1, 26) = 7.32$, $p < 0.012$).

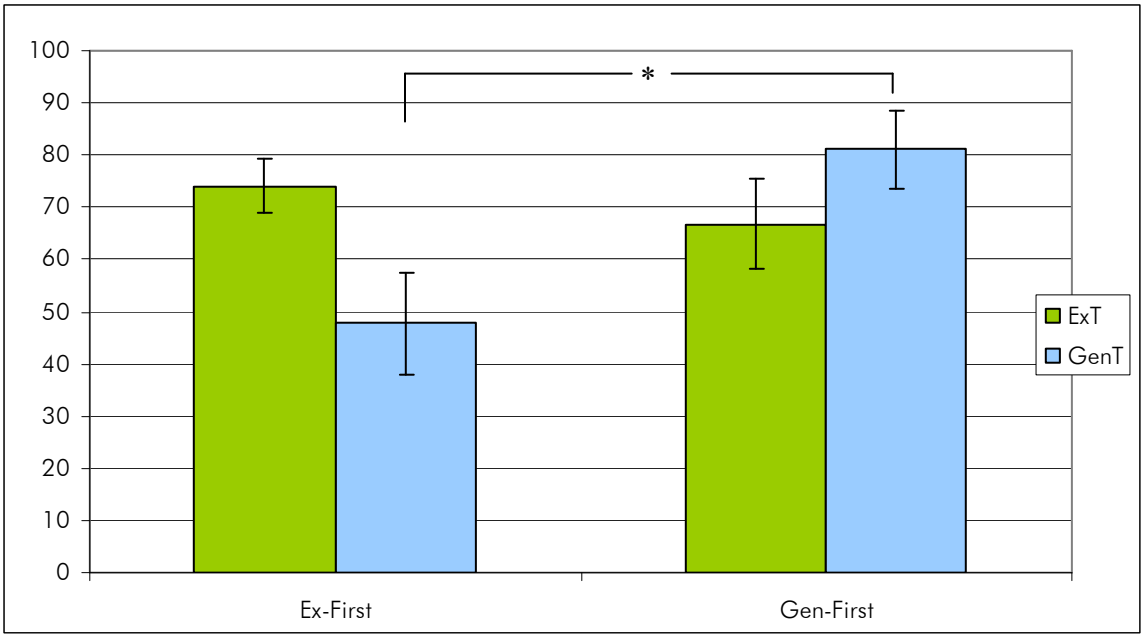


Figure 11. Children's rates of acceptance of test sentences in EXT and GENT by order ** p < 0.05*

As noted, there is no main effect of order on children's rates of acceptance for either context in the interpretation of bare plural subjects. Overall, children do not accept fewer sentences in the EX-FIRST order than in the GEN-FIRST order, or vice versa; however, for both orders, there is a numerical decrease in the rate of acceptance of the sentences in the second block. Children in EX-FIRST accept fewer sentences in GENT in the second block than they do test sentences in EXT in the first block and children in GEN-FIRST accept fewer sentences in EXT in their second block than they do test sentences in GENT in their first block. Recall that the stories were blocked to protect against potential interference effects of the two available interpretations. It now seems as though the first block does in fact have an effect on subjects' rates of acceptance in the second block. In order to look at subjects' unbiased, unprimed rates of acceptance for each interpretation, we can compare the first block from participants in EX-FIRST (i.e., the EXT items) to the first block from participants in GEN-FIRST (i.e., the GENT items).

In block 1, participants in EX-FIRST accepted 74% of the test sentences in EXT and participants in GEN-FIRST accepted 81% of the test sentences in GENT. In block 2, participants in EX-FIRST accepted 46% of the test sentences in GENT and participants in GEN-FIRST accepted 67% of the sentences in EXT. A 2x2 repeated measures ANOVA with order and block as factors shows that there is no main effect of order ($F(1, 26) = 2.88$, $p = 0.102$). However, there is a main effect of block ($F(1, 26) = 6.16$, $p < 0.020$). Participants accepted significantly fewer test sentences in block 2, when they were presented with the second of the two potential interpretations for bare plural subjects. There was no interaction ($F(1, 26) = 0.53$, $p = 0.473$).

The graph below collapses across orders to show the significant difference between block 1 and block 2. Overall, children accepted 76% of test utterances in block 1 and only 57% in block 2. This difference is significant ($t(27) = 2.5$, $p < 0.020$, two-tailed).

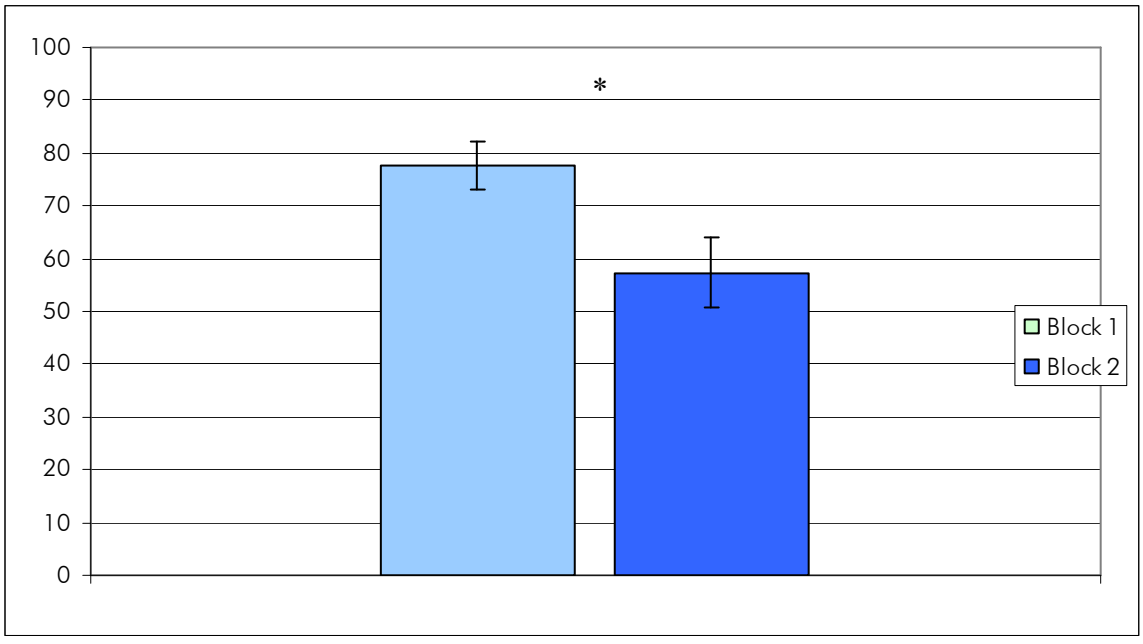


Figure 12. Children's rates of acceptance by block *$p < 0.05$*

The findings reported on in this section show that children have access to the existential interpretation of bare plural subjects. They accept test sentences in the ExT context at an overall rate of 70% (74% of the time in Ex-first and 67% of the time in Gen-first), which is not significantly different from adults. This is the case even though children get no evidence in the input for the availability of this interpretation. This finding supports rejecting the hypothesis that children are restricted to the distributions that are in the input. Instead, these findings support the hypothesis that the MH is part of children's grammars because children accept test sentences in both the ExT and GenT contexts at high rates. This is exactly what is predicted if children have access to both of the interpretations available for bare plural subjects.

It is also important to note that not only do children have access to both interpretations, but the rate at which the test sentences that reflect each interpretation are accepted is not significantly different when there is no interference from the prior context. Figure 13 illustrates the non-significant difference ($t(26) = -0.77$, $p = 0.448$, two-tailed) between the rate of acceptance of test sentences in the ExT context and the GenT context when they occur in block 1. Recall that participants in the two orders hear different contexts in their first block (Ex-first hear ExT sentences and Gen-first hear GenT sentences).

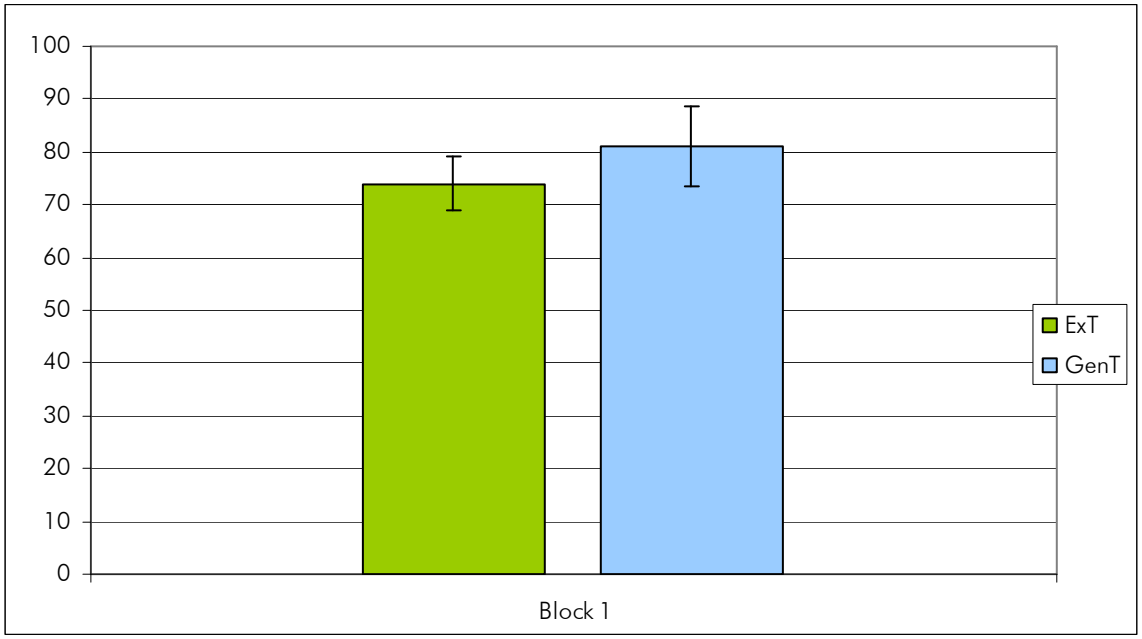


Figure 13. Children's rates of acceptance for EXT and GENT in block 1

3.5.2.2 Adults

Considering now the results from adults alone, recall the preliminary findings reported in Figure 9. Despite predictions that adults should accept both EXT and GENT sentences at high rates, adults accepted the GENT sentences only 32% of the time. While this finding is intriguing on its own, collapsing across the two orders obscures more interesting data.

In EX-FIRST, adults accept 64% of test sentences in EXT and 40% of test sentences in GENT. In GEN-FIRST, they accept 91% of sentences in EXT and only 24% of the test sentences in the GENT context. Single-samples *t*-tests comparing the rate of acceptance of each context to chance shows that the adults in EX-FIRST are at chance for both contexts (though the blocks are significantly different from each other), while the adults in GEN-FIRST are significantly different from chance for both (EX-FIRST, block 1: $t(13) = 1.26$, $p = 0.229$; block 2: $t(13) = -0.096$, $p =$

0.353; GEN-FIRST, block 1: $t(13) = -4.06$, $p < 0.002$; block 2: $t(13) = 7.43$, $p < 0.0001$; all results were two-tailed).

A 2x2 repeated measures ANOVA with order and context as factors reveals that there is no main effect of order ($F(1, 26) = 0.346$, $p = 0.561$), but there is a significant main effect of context ($F(1, 26) = 24.6$, $p < 0.0001$). Adults' rates of acceptance of sentences in the GENT context are significantly lower than their rates of acceptance of test sentences in the EXT context. There is also a significant Order x Context interaction ($F(1, 26) = 5.48$, $p < 0.027$). This is because adults in GEN-FIRST accept significantly more sentences in the EXT context than did adults in EX-FIRST ($t(26) = -2.09$, $p = 0.047$). A correlated samples t-test reveals that the difference in EX-FIRST between the EXT and GENT contexts is non-significant ($t(13) = 1.55$ $p = 0.073$, two-tailed).

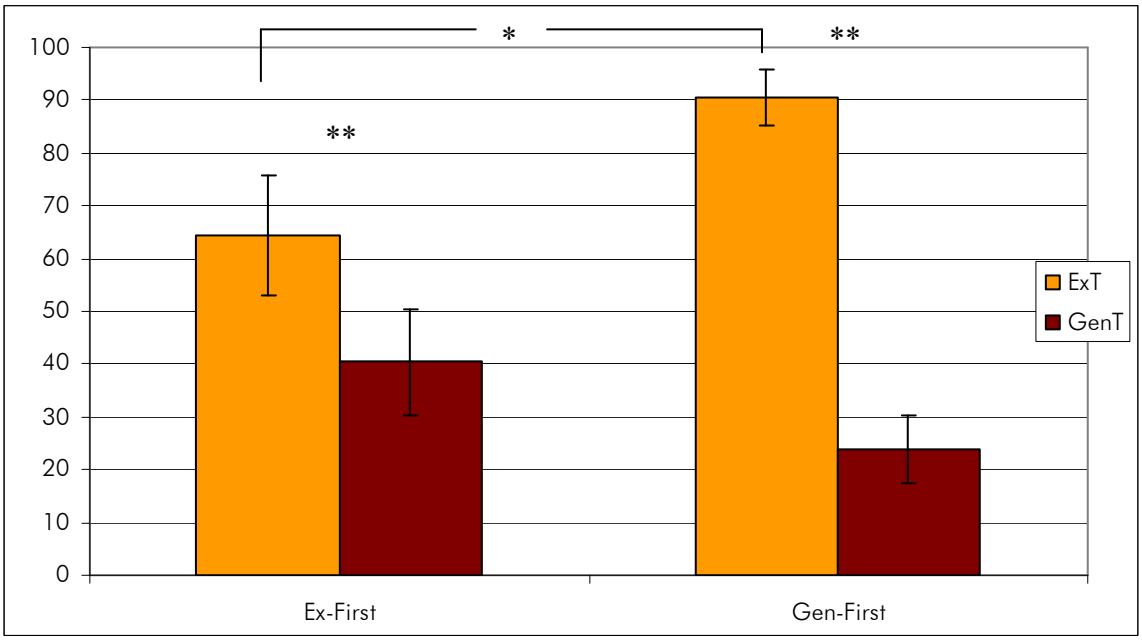


Figure 14. Adults' rates of acceptance for EXT and GENT sentences by order ** $p < 0.05$, ** $p < 0.0001$*

As in the analysis of the child data, it is important to look at the two blocks separately. In order to look at participants' unbiased, unprimed rates of acceptance for each interpretation, we can compare the first block from participants in EX-FIRST (i.e., the ExT context) to the first block from participants in GEN-FIRST (i.e., the GenT context).

The results of block 1 for each of the two contexts appear on the left of Figure 15 and the results of block 2 appear on the right. In block 1, participants in EX-FIRST accepted 64% of the test sentences in ExT and participants in GEN-FIRST accepted 24% the sentences in GenT. In block 2, participants in EX-FIRST accepted 40% of the test sentences in GenT and participants in GEN-FIRST accepted 91% of the test sentences in ExT. The rates of acceptance of the two contexts are significantly different in both blocks. In both blocks, participants accepted significantly more test sentences in the ExT context than they did test sentences in the GenT context (block 1, $F(1, 26) = 9.68$, $p < 0.004$; block 2, $F(1, 26) = 19.4$, $p < 0.0001$).

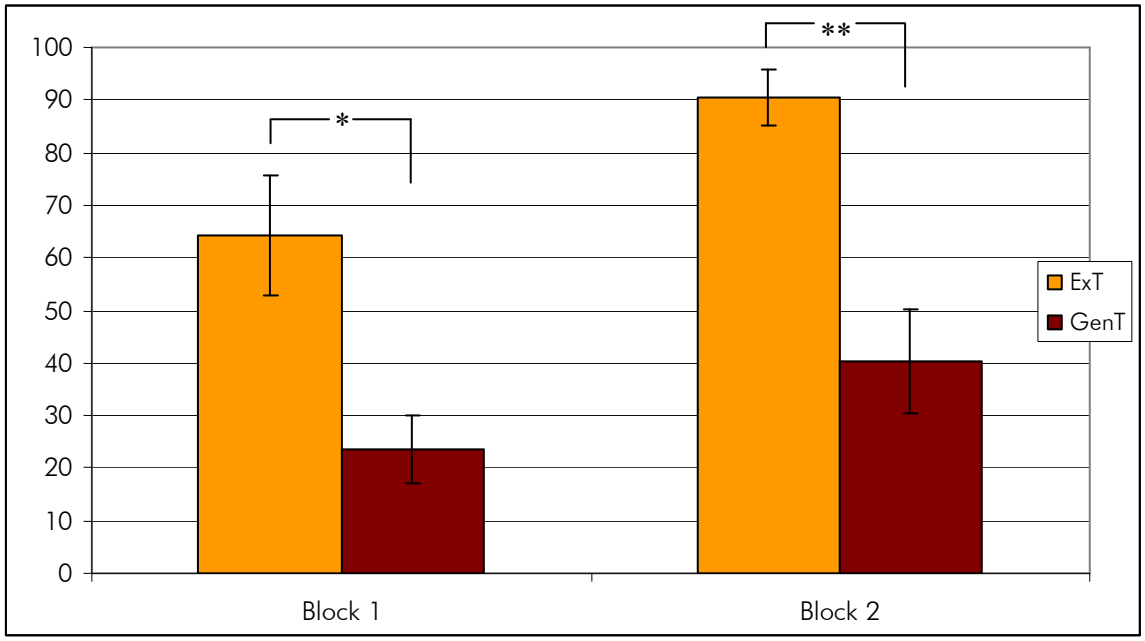


Figure 15. Adults' rates of acceptance for ExT and GenT by block ** $p < 0.004$, ** $p < 0.0001$*

As illustrated in Figure 16, a 2x2 repeated measures ANOVA with order and block as factors shows no main effect of order. However, there is a main effect of block ($F(1, 26) = 5.48$, $p < 0.027$). Overall, participants accepted significantly more test sentences in block 2. There is also an interaction between order and block ($F(1, 26) = 24.6$, $p < 0.0001$) because the participants in EX-FIRST accepted the sentences in their first block (EXT) at higher rates than the sentences they heard in their second block (GENT), while the subjects in GEN-FIRST showed the opposite response pattern, accepting more test sentences in the second block than in the first.

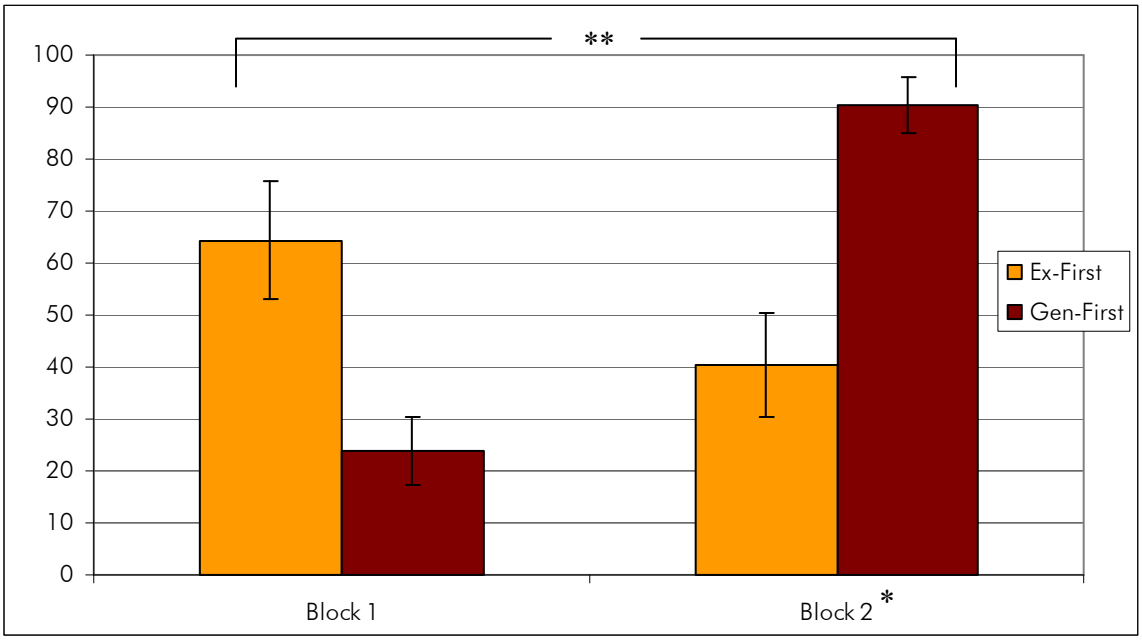


Figure 16. Adults rates of acceptance for Ex-First and Gen-First by block ** $p < 0.05$, ** $p < 0.0001$*

3.6 Discussion

The goal of this experiment was to examine whether the distributional patterns in the input were reflected in children's understanding of bare plural subjects. With respect to children, two hypotheses were explored. The first predicted that if children were limited to using the input to learn about the meanings of indefinite subjects, they would be expected to treat bare plural

subjects as generics and judge test utterances as false when the context makes only the existential interpretation true. If children rejected test sentences with bare plural subjects that were true on their existential interpretation, this would suggest that the input dictates children's interpretation of NPs and that acquisition of the full range of NP meanings must be learned through additional means. Since the MH predicts that children have access to both interpretations of indefinites, such a finding would provide evidence against the MH status as part of children's grammars.

The second hypothesis predicted that if the MH is part of children's grammars, children should have access to both interpretations of indefinite subjects, despite what is in their input. If this is the case, children should behave like adults and accept the test sentences with bare plural subjects both in contexts that make them true on their existential interpretation and in contexts that make them true on their generic interpretation.

The results of this study provide support for the second of these hypotheses, suggesting that the MH is, indeed, part of children's grammars. Children have access to both interpretations and accept the EXT interpretation at the same rate as adults (children, 70%; adults, 77%) collapsing across orders.

A second prediction that was not borne out was that adults would accept test sentences in both EXT and GENT contexts at high rates. Both of these interpretations are logically possible for bare plural subjects; however adult participants show a strong dispreference for the generic interpretation of bare plurals. Examining the speech to children in the corpus studies suggests that adults use bare plural subjects in generic utterances, so this finding is particularly surprising. Recall too that adults were asked to provide justifications for their responses. Without fail, participants who rejected a test sentence with a bare plural subject cited the other interpretation that is available for indefinite subjects in their justification. When adults rejected a *BP*-sentence

in the GENT context, their justification always included a reference to the existential interpretation. When adults rejected a *BP*-sentence in the EXT context, their justifications uniformly included a reference to the generic interpretation, which provides additional evidence that this population of adults has access to the generic interpretation of bare plural subjects. This finding thus raises the question of what features of the materials led to such a strong dispreference for adults to interpret bare plural subjects generically, even though they are produced in speech to children and adults make use of the generic interpretation in providing justifications for rejecting sentences in the EXT experimental context.

One possible explanation is that adults were sensitive to the test sentence form. There were no item effects for adults or children singly, nor were there item effects for the populations as a whole. However, for all the ambiguous *BP*-test sentences, the form of the sentence was bare plural subject + *be* + locative phrase. It's possible that in naturally occurring speech, such utterances are more often interpreted existentially. This type of construction is different from the examples of bare plural subjects that are found in the corpora in Experiments 1 and 2, where indefinite subjects are used to convey information about richly structured categories. In fact, there are no examples of sentences with the form indefinite subject + *be* + locative phrase in the corpus. The only examples of *be* + locative phrase that are similar in form to the test sentences used in this experiment occur with definite subjects.

69) a. The car is on the train. (Adam)

b. The car keys are in my coat. (Tardif E11)

c. All the pictures are in here. (Tardif E22)

Children have had far less input than adults and have had far fewer chances to build a statistical model of the probabilities of utterances of this form. For children, the absence of relevant distributional cues for sentences of this form combined with the relative abundance of generically-interpreted indefinite subjects and the absence of existentially-interpreted indefinite subjects in their input is likely to boost the salience of GENT.

A second possibility is that the generic interpretation was not salient enough for adults. It seems plausible that, for adults, it was difficult to treat the set-up of what was usually true in each story (e.g., as described in §3.3.3, the habitats of the three kinds of animals in the zoo story) to be generically true because it was, in a sense, only true in this particular story. If this is the case, it might be expected that adults would have a strong preference for existentially interpreted bare plural subjects. For children, on the other hand, it may have been easier to limit the domain to the world of the story, to the exclusion of the real world and other possible worlds.

A final point that is important to mention is the effect of order. Both children and adults showed an ordering effect between blocks. What is particularly interesting in this case is that they showed different effects. For children, whatever context they heard in the first block affected their rates of acceptance in the second block by significantly lowering them. In other words, there was a benefit to the interpretation that was previously accessed. In the absence of prior context, both interpretations are highly acceptable to children (74% of test sentences in EXT, 81% of test sentences in GENT), which is exactly what we would expect if bare plural subjects are ambiguous. While both interpretations suppress acceptances of the opposite interpretation in the second block, it is important to note that the existential interpretation has a more persistent effect. Exposure to the EXT context in the first block results in lower second

block acceptance rates (for sentences in GENT) relative to the rate of acceptance in the second block following the GENT context.

The ordering effect for adults, on the surface, seems quite different. First, there is a significant difference for adults in their rates of acceptance of the two contexts in the first blocks. Adults accept sentences in the EXT context 64% and in the GENT context only 24% ($F = 9.68$, $p = 0.004$). This suggests that their unbiased interpretation of utterances with bare plural subjects has a reliable, if not robust, preference for existential subjects, at least in these contexts. The adults in EX-FIRST follow the same general pattern of order effect as the children in EX-FIRST, accepting the test sentence they hear in the first block at a higher rate than the test sentence type they hear in the second block (64% EXT, 40% GENT). However, the difference does not reach significance ($t(13) = 1.55$ $p = 0.073$, one-tailed). In GEN-FIRST, adults accept the sentences in the EXT context in the *second* block at a significantly higher rate than they accept the utterances they hear in GENT in the first block (91% EXT, 24% GENT). In fact, it is only in GEN-FIRST that adults exhibit the projected near-ceiling acceptance rates for even one of the two interpretations.

One possible explanation is that for adults in GEN-FIRST, accessing the EXT interpretation in order to reject the dispreferred interpretation (GENT) boosts the acceptability of EXT when they hear it in the second block. A related possibility is that there is a slight baseline preference for the existential interpretation of bare plural subjects (as revealed in order EX-FIRST, where the first block contains EXT that are interpreted without any interference). If this is the case and adults have a slight preference for existential bare plural subjects, then when adults in GEN-FIRST are presented with the first block of test sentences that are true on the dispreferred interpretation, they are likely to reject them. Because adult participants have already accessed the existential

interpretation of the bare plural subject in the first block, it is easier to access that interpretation in the second block.

Research by Fernanda Ferreira and colleagues (Ferriera et al. 2001) may bear on the issue of the availability of the two interpretations of bare plural subjects and, in fact, unify the findings for children and adults to a certain extent. Ferreira et al. (2001) examined the interpretation and reanalysis of garden path sentences like the example in (70) from their first experiment.

70) While the man hunted the deer ran into the woods.

They found that after reading sentences like this one, participants believed not only that the deer had run into the woods (correctly), but also that the man had hunted the deer (incorrectly). Ferreira et al. conclude that reanalysis of a garden path sentence does not completely erase the initial interpretation. A parse can be revised, but the initial parse is persistent.

For both adults and children, the persistence of a particular interpretation for bare plural subjects seems to influence the acceptability of the other logically possible interpretation. For children, this persistence manifests itself in a preference for the interpretation that the participant heard in the first block. Recall that for both EXT and GENT contexts, children's rates of acceptance of both interpretations in the first block were quite high. Moreover, the acceptance rates were significantly higher in the first block than they were in the second block. The significant difference between the rates of acceptance between block 1 and block 2 could be explained by the persistent activation of the first interpretation that children accessed to accept the test sentences in block 1.

For adults, the persistence of a preferred interpretation manifests itself in a more subtle way. In both orders, there is a preference for the existential interpretation of bare plural subjects. The difference is non-significant in order EX-FIRST, but results in a tendency toward lower rates of acceptance for the GENT context in the second block. This non-significant baseline preference drives the high rates of rejection of the generic interpretation of bare plural sentences in the GENT context of GEN-FIRST. Adults have already accessed the existential interpretation of the bare plural subject in the first block in order to reject the generic interpretation so its persistence in the second block leads to high rates of acceptance in the EXT context.

Chapter 4

4.0 Priming study

In Chapters 2 and 3, I have provided empirical support for analyzing genericity as a property of sentence grammar. I have done so by appealing to Diesing (1992)'s MH. Specifically, I have shown through a series of three experiments that the learner can distinguish indefinites from definites by examining the distribution in the input of different morphological determiners and the syntactic position of NPs. Further, I have shown that the distribution of generics in the input is consistent with the MH, regardless of the fact that it reflects only a subset of the MH's predictions. In Chapter 3, the findings of Experiment 3 have shown that children know about existentially interpreted subjects despite a lack of evidence in the input, supporting my claim that the MH drives learning. However, I have yet to provide evidence that a syntactic explanation for the interpretation of indefinites is the most appropriate analysis. If it is, we would expect to find evidence showing that existentially interpreted indefinite subjects prime other structures requiring VP-internal subjects, while generically interpreted subjects prime other structures requiring subjects interpreted outside of VP.

The present chapter provides further evidence in support of a syntactic explanation for the interpretation of indefinites by showing that *BP*-sentences from Experiment 3 presented in EXT contexts prime other structures requiring VP-internal subjects as compared to the same *BP*-sentences in GENT contexts. In other words, the *BP*-sentences from Experiment 3 can be used to prime interpretations of target sentences that are similar to the primes only because they exhibit similar syntactic representations. In particular, I will test the ability of *BP*-sentences to prime so-

called ISOMORPHIC and NONISOMORPHIC interpretations of quantificational sentences with *every* and *not* as in (71).

71) Every horse didn't jump over the fence.

a. Every_x [[x is a horse] [x ¬ jump over the fence]]

None of the horses jumped over the fence.

b. ¬ Every_x [[x is a horse] [x jump over the fence]]

Not every horse jumped over the fence.

The representation in (71a) represents the isomorphic interpretation of the sentence in (71). Here, *every horse* is interpreted in the surface order of the constituents, outside of the scope of *not*. Note also, that the quantified subject is interpreted in IP and is structurally analogous to generically interpreted indefinite subjects. The representation in (71b) is the nonisomorphic interpretation. Here, *every horse* is interpreted within of the scope of *not*. In this case, the quantified NP (QNP) subject is interpreted in VP and is structurally analogous to existentially interpreted indefinite subjects.

I will show that an existentially or generically interpreted bare plural subject can prime the analogous structural relation for sentences like (71), nonisomorphic or isomorphic, respectively. The finding that generic and existential sentences can be used to prime other quantificational sentences, with completely different meanings, shows that the structural position in the clausal architecture is accessed in interpreting an indefinite subject. This finding provides further support in favor of the MH.

4.0.1 Isomorphism

The terms ISOMORPHIC and NONISOMORPHIC come from Musolino et al. (2000). In a series of studies, these authors found that children tend to interpret NPs in their surface position, interpreting sentences like (71) as (71a), but interpreting sentences like (72) as (72b).

72) The horse didn't jump over every fence.

a. Every_x [[x is a fence] [the horse ¬ jump over x]]

The horse didn't jump over any fence.

b. ¬ Every_x [[x is a fence] [the horse jumped over x]]

The horse didn't jump over every (one of the) fence(s).

Adults easily access both interpretations for both sentences, but children prefer to interpret negation inside the scope of the QNP when the QNP is in subject position but outside the scope of the QNP when the QNP is in object position. However, Musolino at al. could not say whether children's isomorphism was due to linear ordering or syntactic structure.

Lidz & Musolino (2002) addressed this issue comparing English to Kannada, where negation both c-commands and follows object position, unlike English, where it only c-commands object position[64, 65]. They found that it is the scopal relations of two scope-bearing elements (in this case *every* and *not*) that determine how children interpret ambiguous sentences

[64] Lidz & Musolino (2002) also addressed a second problem with Musolino et al. (2000), related to entailment. Because Musolino et al. used the ∀-quantifier in their test sentences, the surface scope entailed the inverse scope. In other words, when the surface scope was true, so was the inverse. Lidz & Musolino's solution to this problem was to use numerically quantified NPs:

i) The horse didn't jump over two fences.
a. There are two (particular) fences that the horse didn't jump over.
b. It's not the case that the horse jumped over two fences (it only jumped over one).

[65] See §1.3.3 for a brief discussion of c-command.

like those in (71) and (72), not simply the surface string. They also found that adults tend to prefer the nonisomorphic interpretation of sentences like (71).

In a series of follow-up studies investigating isomorphism, Musolino & Lidz (2003) found further evidence showing that children and adults show opposite interpretation preferences for sentences like (71) containing negation and QNP. They also found that kids are able to access the nonisomorphic interpretation of such sentences when they are preceded by an affirmative, perhaps because negation is more felicitous in context where it highlights a discrepancy between the actual and expected outcomes (Wason 1965).

73) Every horse jumped over the log and/but every horse didn't jump over the fence.

The findings from the literature on isomorphism are relevant here because the structural relationships involved in interpreting sentences like (71) isomorphically or nonisomorphically are structurally analogous to interpreting bare plural subjects generically and existentially, respectively. Recall that the representation in (71a) represents the isomorphic interpretation of the sentence in (71).

71) Every horse didn't jump over the fence.

a. Every_x [[x is a horse] [x ¬ jump over the fence]]

None of the horses jumped over the fence.

b. ¬ Every_x [[x is a horse] [x jump over the fence]]

Not every horse jumped over the fence.

Here, *every horse* is interpreted outside of the scope of *not* in the surface order of the constituents. The quantified subject NP is interpreted in IP and is structurally analogous to generically interpreted indefinite subjects. The representation in (71b) is the nonisomorphic interpretation. Here, *every horse* is interpreted within of the scope of *not*. In this case, the QNP is interpreted inside VP and is structurally analogous to existentially interpreted indefinite subjects.

If sentences with bare plural subjects in EXT or GENT contexts can prime the analogous structural relation for sentences like (71), nonisomorphic or isomorphic, respectively, this provides further support for the claim that the structural position in the clausal architecture is accessed in interpreting an indefinite subject. This finding would provide further support in favor of the MH.

4.0.2 Priming literature

It is widely known that adults are sensitive to different types of grammatical priming, including the priming of syntactic structures (Bock 1986, Pickering & Branigan 1998, Branigan et al. 2000, Ferriera 2003, inter alia). One analysis of syntactic priming, favored by Branigan et al. (2005), is that effects of syntactic priming arise from a persistent activation of syntactic rules accessed during language production and comprehension. All structures or phrasal categories that allow a particular combinatorial possibility are linked to a particular syntactic rule; when a combinatorial rule is accessed for comprehension (or production) of a particular exemplar, all the related possibilities are activated. Residual activation makes it easier to re-access a structure in subsequent language use.

Until relatively recently, priming studies examined only adult language users. However, a number of recent studies have shown that children are sensitive to priming too. For instance,

using passive sentences, Savage et al. (2003) found that children are sensitive to sentence-level structural priming when there is a considerable degree of overlap between the prime and the target (e.g., when the same noun or pronoun can be used). Huttenlocher et al. 2004 showed that children are sensitive to syntactic priming for the argument structure of verbs. Both of these studies found that unlike adults, preschool-aged children do not transfer syntactic structure across lexical items. The observed effects of priming in these studies were limited to the lexical items used in the primes, revealing a very limited ability to abstract to adult-like representations.

A recent study by Bencini & Valian (in press) found that children are able to abstract away from particular items and make adult-like generalizations. In this study, children in the experimental conditions saw eight pictures depicting fully reversible (active to passive; passive to active) transitive events. They then repeated an active or passive sentence that described the picture (e.g., *The wagon is carrying the presents* or *The presents are carried by the wagon*). These primes were alternated with eight picture descriptions of different events with different inanimate participants (e.g., a knife slicing a lemon). Bencini & Valian found that children in the passive priming condition produced significantly more passives than in the active prime condition (16 passives vs. 3 passives); no passives were produced in a control condition. Unlike previous studies, this particular experiment provides evidence of abstraction because all the "test" items contained new nouns and verbs. Additionally, children's productions were only counted as passives if they were full passives (i.e., SUBJECT was VERBed by OBJECT). Productions did not count as passive if they contained pronouns rather than full subjects.

These findings also show that children exposed to passives on average produced one passive, whereas children exposed to actives produced only 0.2 passives. The magnitude of

differences suggests that a little input for an uncommon structure goes a long way in priming its subsequent use.

Perhaps the most compelling evidence motivating Experiment 4 of this dissertation comes from Viau et al. (2005), which found evidence for LF-priming. Viau et al. (2005) examined the use of structural and pragmatic cues in facilitating access to the nonisomorphic interpretation of sentences like (71). In their first study, children were presented with test stories that showed all three characters in the story successfully completing one task before only two of the three successfully complete a second task. Children in the experimental condition heard three test sentences like (73) followed by three test sentence like (71), whereas children in the control condition heard six test sentences like (71), without explicit mention of the contrast between the two events. Children in the experimental condition did not differ significantly from children in the control condition in the proportion of nonisomorphic responses they provided. This suggests that pragmatic factors alone cannot facilitate access to the nonisomorphic interpretation

For their Experiment 2, Viau et al. (2005) presented children in the experimental condition with a block of three stories where all three characters in the story successfully completed one task before only two of the three successfully completed a second task ("Early Success" (ES)), followed by a block of three stories where all three characters in the story consider a particular task, but in the end decide against it and only two of the three characters do something else ("Early Failure" (EF)). At test, these participants heard test sentences like (71) for both blocks. Participants in the control condition heard six EF stories, where the three characters consider a particular task, but in the end decide against it and only two of the three characters do something else. At test, these participants also heard test sentences like (71). Viau et al. found that in the experimental condition, by training children on three ES stories as described above,

children provided significantly more nonisomorphic responses on the subsequent EF stories. In other words, ES stories can prime nonisomorphic interpretations in EF stories.

Viau et al. (2005) interpret this finding as evidence in favor of LF-priming. Experience with stories that implicitly make the negation in the test sentence felicitous should have no effect on increasing the proportion of nonisomorphic responses to test stories that are less felicitously described with negation, unless the syntactic representation of the test sentence is accessed. Like Bencini & Valian (in press), Viau et al. found that relatively little input for a structure that is uncommon in children's grammars goes a long way in priming its subsequent use.

The findings from these priming studies are relevant to the current study because, first, they show that children are sensitive to syntactic priming, and second that children are able to abstract away from particular lexical items to generalize the syntactic structures they have been primed with. If the analysis of genericity presented in this dissertation is correct, we should see evidence that sentences with bare plural subjects in EXT or GENT contexts can prime the analogous structural relation for universally quantified sentences with negation (nonisomorphic or isomorphic, respectively). To be more precise, bare plural subjects in EXT contexts should have the same priming effect of Viau et al. (2005)'s ES sentences on the interpretation of EF sentences because accepting a bare plural test sentence in the EXT context means that a speaker has activated the VP-internal structure required for the existential interpretation. This structural representation is the same structural representation required for the nonisomorphic interpretation in sentences like (71). EXT contexts represent another way for children to activate the interpretation of the subject within VP, facilitating a VP-internal interpretation of *Every horse didn't jump over the fence*.

4.1 Experiment

This study examined normally developing 4-year-old children growing up in English-speaking homes. The children (n = 30) in this study were recruited from the same two sources as the children who participated in Experiment 3 (Northwestern University's Project on Child Development database and preschools and child care centers in the northern suburbs of Chicago, including Evanston and Northbrook; see §3.1.1 for details) and one additional source, the Center for Young Children at the University of Maryland.

The number of children included in the analyses was 21 (9 girls, 12 boys) ranging in age from 4;0 to 4;9; mean age 4;5. An additional 9 children were run in the study but were excluded from the final analysis for one or more of the following reasons: inability/unwillingness to complete the experiment (n = 2), response bias or innattention during the experiment[66] (n = 6), experimenter error (n = 1).

4.2 Method

The methodology used in this experiment was the same as in Experiment 3. The TVJT was used again, where participants watched a series of stories and heard a puppet's follow-up statement about each story. Participants judged whether or not the puppet had felicitously spoken and they rewarded or punished him accordingly. See §3.2 for more details on the TVJT methodology.

[66] As measured by incorrect responses to filler items.

4.3 Stimuli

4.3.1 Test stories

The test stories in this study consisted of three test stories and two fillers from Experiment 3 and three test stories and two fillers from Viau et al. (2005). The particular stories used in this experiment were chosen at random because there had been no observed item effects in either of the experiments from which the stories were drawn.

Recall that in Experiment 3, all subjects heard the same six test stories (and four fillers). The same versions of the stories from Experiment 3 were used in this study[67]. In Viau et al. (2005), the test stories varied. The three test stories from Viau et al. (2005) that were used in this study were the EF stories. In these stories all three characters in the story consider a particular task, but in the end decide against it and only two of the three characters do something else. All participants head the same test and filler stories balanced for order within their block.

4.3.2 Test sentences

Recall that the goal of Experiment 4 is to test whether bare plural sentences in EXT and GENT contexts can prime nonisomorphic and isomorphic interpretations of sentences with QNPs and negation on the hypothesis that bare plural sentences and QNP – negation sentences have analogous syntactic structures. Thus, the only versions of test sentences from Experiment 3 that were used in this experiment were sentences with a bare plural subject. The between subjects variable was the context of the priming sentence: half the participants heard sentences like (66a) and the other half heard sentences like (66b) as primes. Both examples are repeated here.

[67] A detailed description of a sample test story is provided in §3.3.3. The stories from Experiment 3 that were used here were the "Sister", "Zoo" and "Circus" stories. The fillers were the "Grocery Store" and "Obstacle Course" stories. The reader is reminded that a complete list of the Experiment 3 test items, including story summaries and test sentences is included in Appendix A.

66) a. Crocodiles are in the desert area.

$\exists_x$ x is a crocodile & x is in the desert area

"There are crocodiles in the desert area."

b. Crocodiles are in the river area.

$GEN_{x,t}$ [x is a crocodile & t is a time] x is in the river area at t

"It is a general property of crocodiles that they are in the river area."

The test sentences from Viau et al. (2005) were all of the form in example (71). An actual test sentence is provided here in (74).

74) Every spider didn't hide behind the fence.

All the subjects in Experiment 4 heard a block of three test stories with *BP*-sentences in one of two contexts followed by the same three EF test stories.

The dependent measure was %VP-internal subject interpretations. This was chosen as the dependent measure over %-acceptance responses because it allows for a better comparison of the two blocks. In this study, the sentence types are unrelated semantically, but related structurally. Another way of conceptualizing the dependent variable is to see it as %ExT in block 1 and % nonisomorphic responses in block 2.

4.3.3 Test items

The reader is directed to review §3.3.3 for a discussion of the relationship between the stories and the test sentences for the bare plural sentences; however, it will be helpful to go through one of the stories from Viau et al. (2005) as an example here.

Figure 17. Spiders and Professor Jeff, story begins

As the story about three spiders begins, participants can see three spiders, an owl named "Professor Jeff" a rock, a fence and a tree. The spiders and Professor Jeff are standing on one side of the rock away from the fence and the tree. One of the spiders approaches Professor Jeff and asks if he will play hide-and-seek with the spiders. Professor Jeff agrees, and selflessly volunteers to be "it" first. He goes to the far side of the rock and begins counting. While Professor Jeff is counting, the spiders try to decide where to hide. One spider suggests hiding behind the tree. The other two spiders reject this proposal saying that they always hide behind

the tree and that it's too far away. Then one of the spiders suggests hiding behind the fence instead and liking the idea so much, she goes and hides there. One of the remaining spiders agrees that hiding behind the fence is a good idea and she goes to hide there too. The remaining spider realizes that she needs to hide quickly because Professor Jeff is almost done counting; however, both hiding locations leave something to be desired: the spiders always hide behind the tree and it's far, but she thinks that Professor Jeff will see them through the fence. In the end she decides to stay where she is and cross her legs that Professor Jeff won't find her.

Professor Jeff finishes counting and goes to look for the spiders. First he checks behind the tree, because that's where they always hide. He discovers that they are not there, but notices that they're hiding behind the fence (two of them are). He goes over and "tags" them, announcing that he found them.

Figure 18. Spiders and Professor Jeff, story ends

The puppet then summarizes the story and provides a test sentence like (74), repeated here:

74) Every spider didn't hide behind the fence.

As demonstrated for the stories and test sentences for Experiment 3, the stories and test sentences from Viau et al. (2005) together fulfill two important requirements of the TVJT: plausible dissent and falsification. This pattern of truth conditions for the sentences satisfies the condition of falsifiability for the TVJT because within the context of the story, the test sentence is true on only the nonisomorphic interpretation. Crucially, all of the spiders have considered hiding in multiple locations during the story, inviting participants to consider the possibility that they all hid together in a different location than the one they actually chose. This satisfies the requirement of plausible dissent because if the story had been different in an obvious way that is still consistent with the actual story, it would have been appropriate to reject the claim. For instance, in this particular story, the spiders might have all hidden behind the tree, or they might all have stayed where they were. A complete list of the test items including story summaries and test sentences for the Viau et al. stories used in Experiment 4 is included in Appendix B.

4.3.4 Experimental conditions

As in Experiment 3, the truth conditions of the test sentence were manipulated by varying the test sentence, not the story. In this study, there was only one version of each story and all the participants observed the same six test stories.

Participants were tested with *BP*-sentences and QNP-NEG sentences in one of two orders. In order EX-FIRST, participants heard a block of three test sentences that were true on their

existential interpretation (EXT) followed by a block of three QNP-NEG sentences (n = 11). In order GEN-FIRST participants heard a block of three test sentences that were true on their generic interpretation (GENT) followed by a block of three QNP-NEG sentences (n = 10).

Recall that each version of the two *BP*-sentences is ambiguous, but the context of the story makes only one interpretation of each version true, as described in §3.3.3. If the analysis proposed here is correct, we should expect to see a higher proportion of nonisomorphic interpretations of the QNP-NEG sentences that followed the EXT context than we see following the GENT context.

4.4 Procedure

Children were tested individually in a quiet room. Parents who chose to sit in the room to observe the experiment were instructed to remain quiet and not to say anything about the stories or the child's responses while the experiment was in progress.

4.5 Results

The results of this experiment are displayed in Figure 19. Independent samples t-tests were run to compare participants in the two conditions (EX-FIRST and GEN-FIRST) within each block. The question investigated here is whether participants who were primed with VP-internal subjects accept more nonisomorphic interpretations than participants who were primed with VP-external subjects. The most relevant test of this is the comparison of the two conditions in block 2 but the findings from block 1 were also examined to ensure that participants responded appropriately to the different priming contexts.

Participants in the EX-FIRST condition interpreted 85% of *BP*-sentences existentially in block 1, while participants in the GEN-FIRST condition interpreted only 23% of *BP*-sentences existentially. This difference was significant ($t(19) = 5.6$, $p < 0.0001$, one-tailed[68]), which is precisely what is expected based on the results of Experiment 3 showing that children accept both generic and existential interpretations of *BP*-subjects in the appropriate contexts. If we examine these responses as rates of acceptance (of EXT context for the EX-FIRST participants and of GENT context for the GEN-FIRST participants), participants in EX-FIRST accepted 85% of the *BP*-sentences they heard and participants in GEN-FIRST accepted 77% of the *BP*-sentences they heard, echoing the findings from Experiment 3. The difference, as predicted, is not significant ($t(19) = -0.75$, $p = 0.462$, two-tailed).

In the second block, participants in the EX-FIRST condition interpreted significantly more QNP-NEG sentences nonisomorphically than did participants in the GEN-FIRST condition ($t(19) = 2.11$, $p < 0.0242$, one-tailed[69]). What is somewhat surprising, however, is the number of nonisomorphic responses by participants in the GEN-FIRST condition. They accepted 63% of QNP-NEG test sentences, while they only interpreted 23% of *BP*-subjects inside VP. It would be expected that if participants in the GEN-FIRST condition interpreted the *BP*-test sentences generically, as they did, the rate at which they accepted the nonisomorphic interpretation would be much more attenuated than it actually is.

[68] A one-tailed analysis is justified by the results from Experiment 3 indicating that children interpret BP sentences differently in EXT and GENT contexts. We expect fewer BP-sentences in the GENT context to be interpreted existentially.

[69] A one-tailed analysis is justified by the prediction that the difference between these two conditions should be in a particular direction. The structure of BP sentences in the EXT context should result in more non-isomorphic interpretations of QNP-neg sentences because the two sentence types share a syntactic structure.

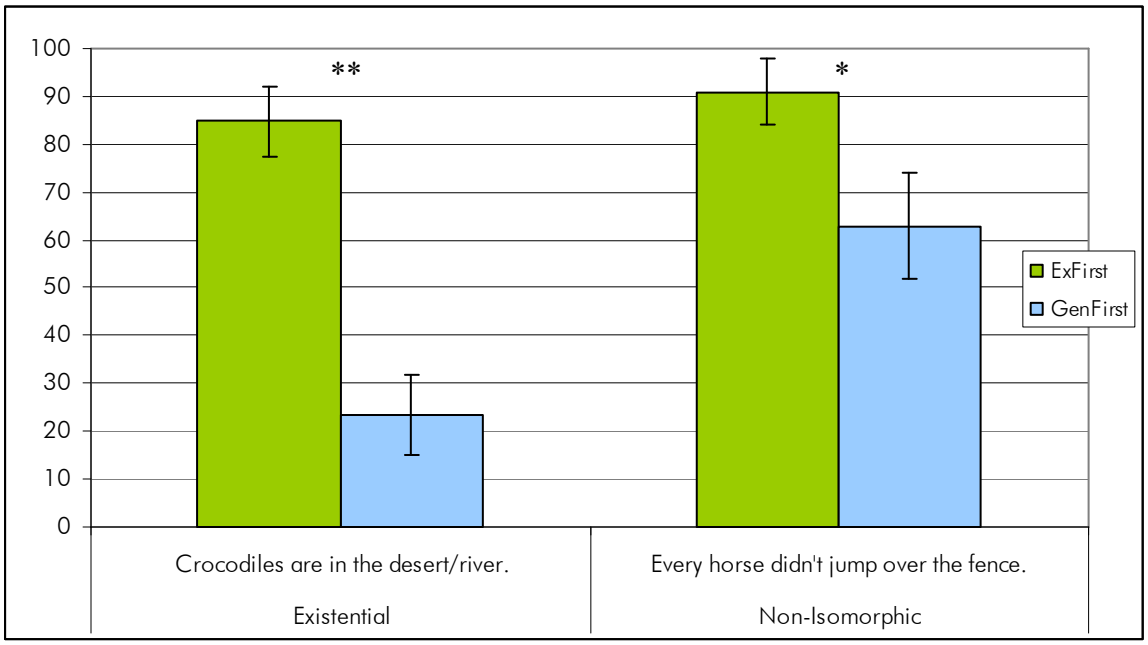


Figure 19. % VP-internal subjects for block 1 and block 2 ** $p < 0.05$, ** $p < 0.001$*

4.6 Discussion

The goal of this experiment was to examine whether ambiguous sentences with bare plural subjects in EXT or GENT contexts can prime the analogous structural relation for sentences like (71), nonisomorphic or isomorphic, respectively. The results of this experiment show that they can, providing additional support for my claim that structural position in the clausal architecture is accessed in interpreting an indefinite subject.

The findings from this study also have broader implications for the syntax-semantics interface. In particular, these findings provide experimental support for the VP-internal subject hypothesis (and the existence of two subject positions) by demonstrating that two semantically unrelated constructions, alike only with respect to having VP-internal subjects, prime each other. These findings also suggest that our analysis of the interpretation of indefinites and universally quantified sentences with negation is right: otherwise we shouldn't see the priming effect.

Finally, these findings provide support for the syntactic nature of LF-priming by demonstrating that interpreting a subject VP-internally in one construction makes it easier to do so in a completely different construction.

One somewhat surprising result was the high percentage of QNP-NEG test sentences that were interpreted nonisomorphically by participants in the GEN-FIRST condition. Subjects in this condition were predicted to interpret more QNP-NEG sentences isomorphically than they did. We might expect that they would have shown results parallel to participants in the EX-FIRST condition interpreting the subjects in QNP-NEG sentences outside of VP at a rate similar to the rate at which they interpreted *BP*-sentences generically. This prediction was not borne out in the data.

One possible explanation is that this population of participants has a default preference for the nonisomorphic interpretation of QNP-NEG sentences and that, in fact, the effect of the GENT context on the interpretation of QNP-NEG sentences was quite robust in lowering their rates of acceptance. In order to test this, the blocks in Experiment 4 should be run in the opposite order. This would provide a baseline rate for isomorphic – nonisomorphic interpretation in this population. If it's the case that these participants do have a preference for the nonisomorphic reading of QNP-NEG sentences, then we might expect to see effects of VP-internal subject priming on the interpretation of *BP*-sentences. Participants would be expected to show high rates of acceptance for *BP*-sentences in the EXT context and lower rates of acceptance (relative to this study) for *BP*-sentences in the GENT context.

Chapter 5

5.0 Conclusion

Generics, in their different varieties, are used to make statements about a category as a whole, not to describe individual members of a category. Generics allow speakers of a language to link disparate individuals, events, or facts by whatever features they are perceived to have in common, even if this represents information that a speaker has never actually observed. Crucially, these perceptions cannot be encoded in the grammar; rather, they vary from speaker to speaker and context to context. Generics are also unique in the way they are used to make generalizations because they seem to refer to the entire category and yet, at the very same time, necessarily *not* the whole category since they allow for exceptions.

All of these properties make generics useful for building representations of categories and therefore for learning about the world. Categories allow speakers to identify things in the world and to draw inferences about individuals based on their membership in a particular category (Carey 1985, Gelman & Markman 1986, Gelman & Markman 1987, Davidson & Gelman 1990, Gelman & Coley 1990, Gelman & Medin 1993, inter alia). Perhaps not surprisingly, generics are very frequent in child-directed speech (Gelman et al. 1998, Pappas & Gelman 1998, Gelman & Tardif 1998) and children themselves produce generics by two years of age (Gelman & Flukes, reported in Gelman 2003: p. 205). But how kids learn generics and the difference between generic and referring language raises a complex question.

One aspect of this question is the problem of generic knowledge (Prasada 2000) and the other is the problem of generic language. The first problem, briefly, is related to how a child is

able to determine whether an observation is applicable to other members of a kind (or only to the individuals she has observed) and if so, how far does the observation extend (i.e. if something is true of tigers, is it true only of tigers, or of zoo animals, or mammals, etc.). Part of possessing generic knowledge involves understanding that an individual is a member of multiple categories. An individual may possesses certain properties that make it a member of a kind, but others that are idiosyncratic.

A more detailed discussion of the first problem is a question for another dissertation[70]. On the other hand, the second problem raised by how kids learn generics and the difference between generic and referring language was precisely the focus.

The goal of this dissertation was to provide a model of how, given the input, a child could acquire the grammar of genericity. To do this, I claimed that genericity is a property of the clausal architecture, not of NPs as the previous literature had suggested. My analysis appealed to a particular linguistic analysis of genericity, Diesing Mapping Hypothesis (Diesing 1992), and made a particular set of assumptions about what is provided to the learner innately. The four studies reported on here were designed to test my claims that combining this particular linguistic analysis with this particular theory of learning would allow children to children acquire the right grammar.

The path that the investigation followed was dictated by the three main predictions the Mapping Hypothesis makes, relevant to acquisition: first, if the learner can find the indefinites in the input, she will know how to interpret them; second, children should have access to all the interpretations that are available for indefinites by virtue of the Mapping Hypothesis being a part

[70] The interested reader is directed to the work of Sandeep Prasada (2000) and Susan Gelman (especially Gelman 2003).

of their grammars, regardless of what may be absent from the input; third, if the grammar of genericity relies on syntactic structure for interpretation, then existentially interpreted indefinite subjects will prime other structures requiring VP-internal subjects, while generically interpreted subjects will prime other structures requiring subjects interpreted outside of VP.

Contrary to previous research, the results reported here show that a learner who possesses a rich system of grammatical representations can determine which NPs in the input are generic and which are not, without being forced to solve the puzzle of a many-to-many mapping between form and function of NPs. This is because the interpretation of an NP relies on its location in the syntactic representation. The child is only required to find the indefinites in the input to make use of the Mapping Hypothesis. The results here show that children can do this by examining the distribution of different morphological determiners and the syntactic position of NPs. These studies also showed that in the input to children, the distribution of indefinite interpretations is a subset of what is predicted by the Mapping Hypothesis, but one that is fully consistent with it.

Children further demonstrate that the Mapping Hypothesis is part of their grammars by showing that they have knowledge of the interpretation of indefinite NPs beyond what is in their input. Children know about the interpretations that are missing from their input. Specifically, children know about existentially interpreted indefinite subjects, which supports the claim that the Mapping Hypothesis drives learning rather than being its output. This also demonstrates the significance of viewing genericity as a property of the syntactic structure, rather than as a property of NP meanings. Children would not be expected to be able to determine the full range of interpretations available unless they had access to the syntactic structures that give rise to the interpretations.

The importance of syntactic structure in the interpretation of indefinite NPs is confirmed by the results showing that existentially interpreted bare plural subjects prime other structures requiring VP-internal subjects, like the nonisomorphic interpretation of sentences with QNPs and negation. Crucially, priming works on sentences that are similar to the EXT primes only because they exhibit similar syntactic representations. Sentences in the GENT context did not show the same degree of priming strength as sentences in the EXT context, but priming with generically interpreted subjects significantly inhibits the interpretation of VP-internal subjects in target sentences.

Taken all together, the ultimate conclusion from the studies presented in this dissertation shows that UG provides children with a rich grammar that they can use to filter the input in meaningful ways. It is thus that they can demonstrate linguistic sophistication beyond what is in their input. They can apply this to the acquisition of generic and referring language, which serve them well as they learn about the world.

REFERENCES

Bacon, J. (1973a). The semantics of generic "the". *Journal of Philisophical Logic*, 2, 323-339.

Bacon, J. (1973b). Do generic descriptions denote? *Mind*, 82, 331-347.

Becker, M. (2000). The acquisition of the English copula. In S.C. Howell, S. A. Fish, & T. Keith-Lucas (eds.), *Proceedings of the Annual Boston University Conference on Language Development*, 24, 104-115.

Bencini, G. M. L. & Valian, V. V. (in press). Abstract sentence representations in 3-year-olds: Evidence from language comprehension and production. *Journal of Memory and Language*.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355-387.

Branigan, H. P., McLean, J. F. & Jones, M. (2005). A blue cat of a cat that is blue? Evidence for abstract syntax in young children's noun phrases. In A. Brugos, M.R. Clark-Cotton & S. Ha (eds.), *Proceedings of the Annual Boston University Conference on Language Development*, 29, 109-121.

Branigan, H.P., Pickering, M.J. & Cleland, A.A. (2000). Syntactic coordination in dialogue. *Cognition*, 75, B13-B25.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Burton-Roberts, N. (1976). On the generic indefinite article. *Language*, 52, 427-448.

Carey, S. (1985). *Conceptual development in childhood*. Cambridge, MA: MIT Press.

Carey, S. & Bartlett, E. (1978). Acquiring a single new word. *Papers & Reports on Child Language Development*, 15, 17-29.

Carlson, G. (1977b). *Reference to kinds in English*. PhD dissertation, University of Massachusetts, Amherst.

Chierchia, G. (1995). Syntactic bootstrapping and the acquisition of noun meanings: The mass/count issue. In B. Lust (ed.), *Syntactic theory and language acquisition* (pp. 301-18). Princeton, NJ: Erlbaum Associates.

Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, 6, 339-405.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Cohen, A. (2001b). On the generic use of indefinite singulars. *Journal of Semantics*, 18, 183-209.

Cohen, A. & Erteshik-Shir, N. (2002). Topic, focus and the interpretation of bare blurals. *Natural Language Semantics*, 10, 125-65.

Contreras, H. (1987). Small clauses in English and Spanish. *Natural Language and Linguistic Theory*, 5, 225-243.

Crain, S. & McKee, C. (1985). The acquisition of structural restrictions on anaphora. In S. Berman, J.W. Choe, & J. McDonough (eds.), *Proceedings of NELS*, 16, 94-110.

Crain, S. & Thornton, R. (1998). *Investigations into Universal Grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge: MIT Press.

Dahl, Ö. (1975). On generics. In E. Keenan, (ed.). *Formal semantics of natural language* (pp. 99-111). Cambridge: Cambridge University Press.

Davidson, D. (1967). The logical form of action sentences. In N. Rescher (ed.), *The logic of decision and action* (pp. 81-95). Pittsburgh, PA: University of Pittsburgh Press.

Davidson, N. & Gelman, S. (1990). Inductions from novel categories: The role of language and conceptual structure. *Cognitive Development*, 5, 151-176.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research*, 30, 3-20.

Ferreira, V. S. (2003). The persistence of optional complementizer mention: Why saying a "that" is not saying "that" at all. *Journal of Memory and Language*, 48, 379-398.

Fisher, C. & Tokura, H. (1995). The Given-New Contract in Speech to Infants. *Journal of Memory and Language*, 34, 287-310.

Fritzley, V.H., & Lee, K. (2003). Do young children always say yes to yes–no questions? A metadevelopmental study of the affirmation bias. *Child Development*, 74, 1297-1313.

Gelman, S. (1988). The development within natural kind and artifact categories. *Cognitive Psychology*, 20, 65-96.

Gelman, S. (2003). *The essential child*. New York: Oxford University Press.

Gelman, S. (2004). Learning words for kinds. In D. Hall & S. Waxman (eds.), *Weaving a lexicon* (pp. 445-84) Cambridge, MA: MIT Press.

Gelman, S. & Coley, J. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old-children. *Developmental Psychology*, 26, 796-804.

Gelman, S. & Markman, E. (1986). Categories and induction in young children. *Cognition*, 23, 183-209.

Gelman, S. & Markman, E. (1987). Young children's inductions from natural kinds: the role of categories and appearances. *Child Development*, 58, 1532-1541.

Gelman, S. & Medin, D. (1993). What's so essential about essentialism? A different perspective on the interaction of perception, language and conceptual knowledge. *Cognitive Development*, 8*,* 157-167.

Gelman, S. & Raman, L. (2003). Preschool children use linguistic form class and pragmatic cues to interpret generics. *Child Development, 74,* 308-25.

Gelman, S., Star, J. & Flukes, J. (2002). Children's use of generics in inductive inferences. *Journal of Cognition and Development*, 3, 179-199.

Gelman, S. & Tardif, T. (1998). Generic noun phrases in English and Mandarin: An examination of child-directed speech. *Cognition*, 66, 215-48.

Gelman, S., Coley, J., Rosengren, K., Hartman, E. & Pappas, A. (1998). Beyond labeling: The role of maternal input in the acquisition of richly structured categories. *Monographs of the Society for Research in Child Development*, 63 (serial no. 253).

Graham, S., Cameron, C. & Welder, A. (2005). Preschoolers' extension of familiar adjectives. *Journal of Experimental Child Psychology*, 91, 205-226.

Graham, S., Welder, A. & McCrimmon, A. (2003). Hot dogs and zavy cats: Preschoolers' and adults' expectations about familiar and novel adjectives. *Brain and Language*, 84, 16-37.

Heibeck, T. & Markman, E. (1987). Word learning in children: An examination of fast mapping. *Child Development*, 58, *4,* 1021-1034.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases*, PhD dissertation, University of Massachusetts, Amherst.

Hollander, M., Gelman, S. & Star, J. (2002). Children's interpretation of generic noun phrases. *Developmental Psychology*, 38, 883-94.

Huttenlocher J., Vasilyeva, M. & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, 50, 182-195.

Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, & M. Stokhof (eds.), *Formal methods in the study of language: Proceedings of the third Amsterdam colloquium* (pp. 277-322). Amsterdam: Mathematical Centre.

Keil, F. (1992). *Concepts, kinds and cognitive development*. Cambridge: MIT Press. Available from http://cognet.mit.edu. (Reprinted from *Concepts, kinds and cognitive development*. Cambridge: MIT Press, 1989)

Kitagawa, Y. (1994). *Subjects in Japanese and English*. Outstanding dissertations in linguistics. New York, NY: Garland Press.

Kleiber, G. (1989a). LE générique: un massif? *Langagues*, 94, 73-113.

Koopman, H. & Sportiche, D. (1991). The position of subjects. *Lingua*, 85, 211-258.

Kratzer, A. (1995). S-level and I-level predicates. In G. Carlson & F.J. Pelletier (eds.), *The generic book* (pp. 125-175). Chicago: The University of Chicago Press.

Kratzer, A. (1998). Scope or Pseudoscope? Are there Wide-scope Indefinites? In Susan Rothstein (ed.), *Events in Grammar* (pp. 163-196). Dordrecht: Kluwer.

Krifka, M., Pelletier, F. J., Carlson, G. N., ter Meulen, A., Link, G. & Chierchia, G. (1995). Genericity: An introduction. In G. Carlson & F.J. Pelletier (eds.), *The generic book* (pp. 1-124). Chicago: The University of Chicago Press.

Kuroda, S. (1992). Whether we agree or not: A comparative syntax of English and Japanese. In S. Kuroda, *Japanese syntax and semantics: Collected papers* (pp. 315-357). Dordrecht: Kluwer. (Reprinted from *Linguisticae Investigationes*, 12, pp. 1-47, 1988)

Lawler, J. (1972). Generic to a Fault. In P. Pervanteau, J. Levi & G. Phares (eds.), *Proceedings of CLS*, 8, 247-258.

Levin, B. (1993) *English verb classes and alternations: A preliminary investigation* (pp. 188-196). Chicago: University of Chicago Press.

Lewis, D. (1975). Adverbs of quantification. In E. Keenan (ed.), *Formal semantics of natural language* (pp. 3-15). Cambridge: Cambridge University Press.

Lidz, J. & Musolino, J. (2002) Children's command of quantification. *Cognition*, 84, 113-154.

Lyons, J. (1977). *Semantics*. New York: Cambridge University Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*, Third edition. Mahwah, NJ: Lawrence Erlbaum Associates.Matthewson, L. (2004)

Matthewson, L. (2004). Strategies of quantification: Evidence from the world. Paper presented at the Strategies of Quantification Conference, York, England. To appear in K. Gil and G. Tsoulas (eds.).

May, R. (1985). *Logical form: Its structure and derivation*. Cambridge, MA: MIT Press.

May, R. (1991). Syntax, semantics and logical form. In A. Kasher (ed.), *The Chomskyan turn* (pp. 334-359). Cambridge, MA: Blackwell.

McCloskey, J. (1996). On the scope of verb movement in Irish. *Natural Language and Linguistic Theory*, 14, 47-104.

Milsark, G. (1974). Existential sentences in English. Ph.D. dissertation, MIT.

Musolino, J. & Lidz, J. (2003). The scope of isomorphism: turning adults into children. *Language Acquisition*, 11, 277-291.

Musolino, J., Crain, S. & Thornton, R. (2000). Navigating negative semantic space. *Linguistics*, 38, 1-32.

Ojeda, A. (1991). Definite descriptions and definite generics. *Linguistics and Philosophy*, 14, 367-397.

Pappas, A. & Gelman, S. (1998). Generic noun phrases in mother-child conversations. *Journal of Child Language*, 25, 19-33.

Partee, B. (1986). Noun-phrase interpretation and type-shifting principles. In J. Groenendijk, D. de Jongh & M. Stockhoff (eds.). *Studies in discourse representation theory and the theory of generalized quantifiers*, volume 8 of the Groningen-Amsterdam Studies in Semantics. Dordrecht: Reidel

Partee, B. (1992) Topic, focus and quantification. In A. Wyner & S. Moore (eds.), *Proceedings of First Annual Conference on Semantics and Linguistic Theory*, 1, 159-187.

Partee, B. (1995). Quantificational Structures and Compositionality. In E. Bach, E. Jelinek, A. Kratzer, & B. H. Partee (eds.), *Quantification in natural languages* (pp. 541-601). Dordrecht: Kluwer.

Pesetsky, D. (1989). Language-particular processes and the earliness principle. Unpublished manuscript. Retrieved April 27, 2007 from http://web.mit.edu/linguistics/people/faculty/pesetsky/publications.html.

Pickering, M. J. & Branigan, H. P. (1998). The representation of verbs: evidence from syntactic priming in language production. *Journal of Memory and Language*, 39, 633-651.

Prasada, S. 2000. Acquiring generic knowledge. *Trends in Cognitive Sciences*, 4, 66- 72.

Prince, E. (1992). The ZPG letter: Subjects, definiteness, and information-status. In S. Thompson & W. Mann (eds.), *Discourse description: Diverse linguistic analyses of a fund-raising text* (pp. 295-325). Philadelphia/Amsterdam: John Benjamins.

Rooth, M. (1995). Indefinites, Adverbs of Quantification and Focus Semantics. In G. Carlson & F.J. Pelletier (eds.), *The generic book* (pp. 265-99). Chicago: The University of Chicago Press.

Savage, C., Lieven, E., Theakston, A. & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, 6, 557-567.

Savage, C., Lieven, E., Theakston, A. & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development*, 2, 27-49.

Schubert, L. K. & Pelletier, F. J. (1989). Generically speaking; Or using discourse representation theory to interpret generics. In G. Chierchia, B. Partee & R. Turner (eds.), *Properties, types and meanings*, vol. 2 (pp. 193-268). Dordrecht: Kluwer.

Spears, A. K. (1974). On the notion occasion and the analysis of aspect. In M. W. La galy, R. A. Fox & A. Bruck (eds.), *Proceedings of CLS*, 10, 672-683.

Sportiche, D. (1988). A Theory of Floating Quantifiers and its Corollaries for Constituent Structure. *Linguistic Inquiry*, 19, 425-450.

Viau, J., Lidz, J. & Musolino, J. (2005). Syntactic priming of LF representations in 4-year-olds. Poster presented at *BUCLD 30*.

von Fintel, K. & Matthewson, L. (in press). Universals in semantics. *The Linguistic Review*.

Wason, P. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavio*r, 4:7-11.

APPENDIX A

Test items for Experiment 3

Alice at the zoo

The zookeeper tells Alice where all the different animals live: the monkeys in the jungle, the crocodiles in the river area, the zebras in the savanna. The zookeeper goes off to work and Alice decides to go see the monkeys. While Alice is over by the jungle area looking at the monkeys, the zookeeper notices that the river where the crocodiles live is filthy and he needs to clean it right away. He has to move the crocodiles to do it, so he takes the crocodiles from the river habitat they normally live in and puts them in the desert area. He then goes back to the river to clean.

Alice decides to go see the zebras in the savanna next and after watching them for a while decides she will go see the crocodiles in the river area. To her surprise, when she arrives at the river area, the crocodiles are gone and the zookeeper is cleaning. He tells her that the river was dirty and that he had to move the crocodiles in order to clean it. She walks away, dejected, but perks up as she notices that that there's something in the desert area. She sees that it's the crocodiles and is happy to see them even if she does not get to see them swim.

Hmm. That was a story about a girl who went to Mr. Sam's zoo. I think I know something about the zoo …

BP	Crocodiles are in the desert area.	ExT = True
	Crocodiles are in the river area.	GenT = True
There	There are crocodiles in the desert area.	ExT = True
	There are crocodiles in the river area.	GenT = False
Usually	Crocodiles are usually in the desert area.	ExT = False
	Crocodiles are usually in the river area.	GenT = True

Max at the toystore

Max arrives at the toystore and meets Eric, who is the owner of the toystore. He shows Max a map of the store and tells him where he can find all the different toys: the skateboards on the first floor, the planes on the second floor, the horses on the third floor. Max decides to start on the first floor with the skateboards. He really likes them and thinks he wants to get them, so, since there are only three left, he takes them with him and goes to look at the toys on the other floors.

Max considers getting some planes on the second floor, but ultimately decides to get some horses up on the third floor. He puts the skateboards down on the third floor and goes to buy the horses.

Hmm. That was a story about a boy who went to a toy store. I think I know something about the store …

BP	Skateboards are on the third floor.	ExT = True
	Skateboards are on the first floor.	GenT = True
There	There are skateboards on the third floor.	ExT = True
	There are skateboards on the first floor.	GenT = False
Usually	Skateboards are usually on the third floor	ExT = False
	Skateboards are usually on the first floor.	GenT = True

Talking dragons at the circus

Two girls go to the circus. They find a sign outside the tent telling them where they can see all the different acts: the lions in the yellow ring, the acrobats in the blue ring, the talking dragons in the green ring. They decide to go see the lion tamer in the yellow ring first. They really enjoy the act, but get bored and decide to go see the acrobats in the blue ring. Meanwhile, the talking dragons in the green ring decide that they are tired and need to take a break. The dragons put up a sign indicating that they've gone outside on break and leave the ring empty.

After watching the acrobats for a while, the girls decide to go see the talking dragons. When they get to the green ring, they notice that the dragons are nowhere to be found; however, they see the sign that the dragons put up to indicate that they went outside on break. The girls debate whether to wait around or come back tomorrow to see the dragons.

Hmm. That was a story about the three-ring circus. I think I know something about the circus…

BP	Talking dragons are outside.	ExT = True
	Talking dragons are in the green ring.	GenT = True
There	There are talking dragons outside.	ExT = True
	There are talking dragons in the green ring.	GenT = False
Usually	Talking dragons are usually outside.	ExT = False
	Talking dragons are usually in the green ring.	GenT = True

Playing at the park

Three friends have had a fun day playing at the park. All the toys they played with are spread out in front of them. Two of the children have to go home for dinner right away and can't help clean up the mess, so one of the children is left to put everything away by himself. Luckily, as he's about to start, the park ranger shows up to remind him of where all the different types of toys belong: the soccer balls in the box by the soccer field, the basketballs in the shed by the

basketball court, the baseballs in the bag by the baseball diamond. The park ranger has to leave too, but she offers to help the boy tomorrow if he has any trouble.

First, the boy puts away the baseballs in the ball bag by the baseball diamond. Then he puts the basketballs in the shed by the basketball court. Finally, he takes the soccer balls over to the soccer field. He tries to put them in the appropriate box, but it's locked and he doesn't have a key. Rather than leave the soccer balls loose, the boy decides to put them in the shed with the basketballs and come back the following day to tell the park ranger.

Hmm. That was a story about a boy cleaning up at the park. I think I know something about the park…

BP	Soccer balls are in the shed.	ExT = True
	Soccer balls are in the box.	GenT = True
There	There are soccer balls in the shed.	ExT = True
	There are soccer balls in the box.	GenT = False
Usually	Soccer balls are usually in the shed.	ExT = False
	Soccer balls are usually in the box.	GenT = True

Ali in her sister's room

Megan is getting ready to go to school and her little sister Ali is eating her breakfast. Megan likes to be very tidy, so she cleans up her room before she goes to school, putting all her things away: the hats in the closet, the pillows under the covers of her neatly made bed, the toys in the toy chest. Just as she's about to leave, she realizes that while she's gone, Ali is going to play with her toys and make a mess in her room. To prevent this, she decides to hide her toys in a different location. Instead of keeping them in the toy chest, she hides them behind her TV and then goes to school.

Sure enough, as soon as Megan is gone, Ali decides to play with Megan's things. She decides to play with the toys, but when she goes to the toy chest to get them, she can't find them. She's disappointed and decides instead to just watch TV.

Hmm. That was a story about a girl playing in her sister's room. I think I know something about the room…

BP	Toys are behind the TV.	ExT = True
	Toys are in the toy chest.	GenT = True
There	There are toys behind the TV.	ExT = True
	There are toys in the toy chest.	GenT = False
Usually	Toys are usually behind the TV.	ExT = False
	Toys are usually in the toy chest.	GenT = True

Ariel and the circus train

Ariel loves trains so she goes to the train yard to go on a tour. The train conductor asks Ariel if she'd like to have a tour of a train that's taking some animals to the zoo so she can see where the animals ride. Ariel says she would, so the conductor tells her where the different animals ride: the cats in the engine, the elephant in the middle car, the chickens in the caboose. Ariel decides to see the elephant first because she's never seen one up close. As they're going to look at the elephant, the conductor reminds her that as they go from car to car, she needs to be sure to shut the door to the car – he doesn't want the wrong animals getting in the wrong cars.

After they look at the elephant, Ariel decides she wants to see the chickens in the caboose. She forgets the conductor's warning and leaves the door open when they leave. On their way up to the engine to see the cats, Ariel and the conductor stop one to see the elephant one more time. While they're looking at the elephant, the cats in the engine decide that they're bored and that they'd like to go scare the chickens in the caboose for a little fun. They sneak back to the caboose and jump in and scare the chickens. When Ariel and the conductor arrive at the engine, the cats aren't there, so they decide to go look in the other cars. They find them in the caboose.

Hmm. That was a story about a girl going on a tour of a train. I think I know something about the train …

BP	Cats are in the caboose. Cats are in the engine.	ExT = True GenT = True
There	There are cats in the caboose. There are cats in the engine.	ExT = True GenT = False
Usually	Cats are usually in the caboose. Cats are in the engine.	ExT = False GenT = True

Fillers for Experiment 3

Elmo at the grocery store

Elmo and Cookie Monster have just come back from a hike and Cookie wants to buy Elmo a treat. He tells Elmo where everything is: the popsicles are in the freezer case, cookies are in the cookie aisle, the cupcakes are in the bakery. Elmo decides to go look at the popsicles first and they look so good that he decides to take them because there's only one box left. However, he can't resist looking at the cookies. The popsicles seem better, but there are still the cupcakes in the bakery and they look best of all. Elmo decides to get a cupcake instead of the popsicles. He puts the popsicles down in the bakery and goes to find Cookie Monster.

Hmm. That was a story about Elmo and Cookie Monster at the grocery store. I think I know something about the store.

Elmo put the popsicles in the bakery. = True

Elmo left the popsicles in the freezer section. = False

Nathan at the garden

Nathan and his dad are planting vegetables and flowers in the community garden. Nathan's dad is tired and wants to rest, so he tells Nathan where to put everything: the tomatoes go next to the hose, the flowers go in the sun next to the fence, the peppers in the shade next to the tree. While Nathan's dad is resting, Nathan gets to work. First he plants the tomatoes next to the hose, then the peppers in the shade next to the tree. However, when it comes time to plant the flowers, Nathan decides that the purple flowers would look really nice next to the yellow peppers and disregards his father's instructions, putting the flowers in the shade next to the tree (like the peppers) and not in the sun next to the fence.

Hmm. That was a story about a boy named Nathan who worked at the community garden with his dad. I think I know something about the garden…

Nathan planted the flowers in the shade. = True

Nathan planted the flowers in the sun. = False

Concert

At the school concert, a girl named Belle is in charge of taking everyone to their seats. Her teacher Mr. Hodges reminds her where everyone is supposed to sit: the parents sit in the chairs in the middle, the kids sit in the grass, the dogs sit over in the dirt. As the guests arrive, Belle takes them to their assigned areas.

Hmm. That was a story about a girl named Annie who took people to their seats at a school concert. I think I know something about the concert…

Belle took the children to the grass. = True

Belle put the dogs in the grass. = False

Obstacle course

Pound Puppy, SpongeBob and Care Bear are standing around trying to think of something fun to do. Pound Puppy gets the idea to race around an obstacle course. SpongeBob agrees to race, but the Care Bear says he'll be the judge of who's the fastest. Pound Puppy goes first and she goes really fast on a scooter. SpongeBob goes next, but his blue truck breaks down and he goes very slowly. Care Bear announces that Pound Puppy was the winner.

Hmm. That was a story about some friends racing around at an obstacle course. I think I know something about it…

The scooter was the fastest. = True

The blue truck was the fastest. = False

APPENDIX B

Test items for Experiment 4

Block 1
Test items from Experiment 3
Alice at the zoo, Max at the toy store, Talking dragons at the circus

Fillers from Experiment 3
Elmo at the grocery store, Obstacle course

Block 2
Early Failure sentences from Viau et al. 2005

Girls
Three girls decide to catch stuff. They consider catching snakes, but decide not to because snakes are gross and slimy. Two of the girls then decide to catch starfish. The third girl is afraid of the water where the starfish are and won't go in.

Hmm. That was a story about some girls who wanted to catch stuff. I think I know something that happened…

Every girl didn't catch a starfish. Nonisomorphic = True, Isomorphic = False

Spiders
Three spiders decide to play hide-and-seek with Professor Jeff. They consider hiding behind the tree but decide not to because the tree is too far away and they always hide there. Then two of the spiders decide to hide behind the fence. The third spider is afraid of being caught and stays put. Sure enough, Professor Jeff finds the two spiders who hid behind the fence.

Hmm. That was a story about some spiders who played hide-and-seek with Professor Jeff. I think I know something that happened…

Every spider didn't hide behind the fence. Nonisomorphic = True, Isomorphic = False

Butterflies
Three butterflies decide to do some flying. They consider flying to the forest but decide not to because it looks boring. Then two butterflies decide to fly to the city for some great views. The third butterfly decides to stay where he is because it's cool and quiet.

Hmm. That was a story about some butterflies who wanted to do some flying. I think I know something that happened…

Every butterfly didn't go to the city. Nonisomorphic = True, Isomorphic = False

Fillers from Viau et al. 2005

Whales
Three whales relax at the beach. They all play with beach balls. Two swim home for dinner after playing, but the third stays on the beach to play with his ball some more.

Every whale played with a ball. = True

Every whale swam home for dinner. = False

Dinosaurs
Three kid dinosaurs decide to do something nice for their teacher Rex, whose birthday is today. They consider drawing pictures for Rex, but their pencils don't have erasers and they're afraid of making mistakes. Instead, they decide to sing "Happy Birthday" to Rex.

The dinosaurs didn't draw pictures. = True

The dinosaurs didn't sing "Happy Birthday" = False