Semantic Manifestations of the Developing Theory of Mind

Natalia Rakhlin, PhD
University of Connecticut, 2007

I investigate a connection between ways in which children typically interpret quantified sentences and an aspect of cognition known as the Theory of Mind or ToM. I have looked at children's errors known as quantifier-spreading (q-spreading) and argued that these errors, previously regarded as semantic, can be analyzed as a consequence of children's immature ToM, whereby a weakness in their ability to infer what is relevant or salient to others causes children to misinterpret quantified sentences, semantics of which requires integrating contextual information. Once children's difficulties with reading vague contextual clues are taken into account, we can see that their knowledge of semantics of quantifiers is adult-like as early as can be tested.

The effect of developing ToM on interpreting quantified sentences becomes visible during selection of values for covert quantifier domain restriction variables, particularly in the interpretation of indefinites. I hypothesize that visual asymmetry triggers "q-spreading" responses not because of children's deviant semantic form, in which the universal ranges over both the agent and theme NPs and consequently requires the denotations of the two to be 'symmetrical' (i.e. members of the agent-set to be exhaustively paired with those of the theme-set and vice versa). Instead, I suggest that contextual asymmetry (an unpaired individual with the property denoted by the indefinite) causes "q-spreading" because children use the asymmetry in the visual context as a pragmatic clue when deciding how to select domain restrictions for the indefinite. Thus, if the picture contains a number of boy/wagon pairs and a salient unpaired wagon (in case of the sentence "every boy is pulling a wagon"), children are likely to see its salience as an indication of its relevance for the speaker and construct the domain restriction for the indefinite to include exactly one object – the extra wagon). Consequently, the sentence receives the "wide scope" indefinite interpretation, false in the situation depicted in this scenario. This theory leads to certain predictions with respect to quantifier scope interactions, relationship between visual salience and error rate, and non-universally quantified sentences triggering q-spreading-like errors, which I test in the experiment reported in the dissertation.

Semantic Manifestations of the Developing Theory of Mind

Natalia Rakhlin

M.A., University of Montana, 1994
M.A., University of Connecticut, 2004

A Dissertation
Submitted in Partial Fulfillment of the Requirements
for the degree of Doctor of Philosophy
at the University of Connecticut
2007

2007

APPROVAL PAGE

Doctor of Philosophy Dissertation

Semantic Manifestations of the Developing Theory of Mind

Presented by
Natalia Rakhlin, B.A., M.A.

Major Advisor _____
Dr. Diane Lillo-Martin

Major Co-Advisor _____
Dr. William Snyder

Associate Advisor _____
Dr. Yael Sharvit

University of Connecticut
2007

Acknowledgments

I would like to express my gratitude to the people without whom this project would never have been completed. First and foremost, I want to thank my advising committee, Diane Lillo-Martin, William Snyder and Yael Sharvit, for the incredible generosity with their time and advise, from which I have greatly profited. I am very thankful to them for giving me space and encouragement to explore seemingly eccentric ideas and for recognizing their potential value. I am even more grateful for their critical judgment and an ability to see logical and empirical flaws with great clarity and for insisting on never lowering the bar. I am especially indebted to Diane Lillo-Martin, who not only taught my first language acquisition course and helped me run my first language acquisition experiment, but weekly conversations with whom were my only regular contact with the 'outside world' during the long and lonely process of writing this dissertation and at times my sole motivation to get the work done.

I want to thank Tom Roeper, Bart Hollenbrandse and the UMass-Smith acquisition community for their interest in my work and helpful suggestions. The same goes to the linguistic community at the University of Connecticut. I am grateful to all for their friendship and support over the years.

Finally, I have to thank my family – Bruce and Daniela. I cherish your patience, willingness to listen and an ability to do so without being bored to death, and most of all for believing in me.

Chapter 1

Introduction

1.1 Goal

In this study I will study the nature of certain well-documented interpretive errors pre-school children make with quantified sentences cross-linguistically in order to gain a better understanding of the relationship between semantics and pragmatics in children's knowledge of quantification. My goal in pursuing this question is to contribute to one of the central questions in the theory of language acquisition, namely a problem of how to resolve the tension arising from two seemingly contradictory lines in the current language acquisition research. On the one hand, beginning with Chomsky's famous "as-if-instantaneous" model of acquisition (Chomsky 1965), the generative theory of language acquisition has developed around the core theoretical assumption of children's innate grammatical competence (Universal Grammar or UG), all major subsequent theoretical developments having been further articulating or refining the original hypothesis (e.g. Strong Continuity Assumption (Pinker 1984), Principles and Parameters framework (Chomsky 1981), Lexical Parameterization hypothesis (Borer 1984), Very Early Parameter Setting (Wexler 1995) etc.). On the other hand, a substantial body of data has been collected to show linguistic behavior in children that seems to be arbitrarily different from that of adults prompting some researchers to argue that certain aspects of UG do not become available to the child from the start, but develop on a maturational schedule (e.g. Borer and Wexler 1987, 1992, Philip 1995, Wexler, K. 2003, Drozd 2001, Musolino 2003). The important question I have alluded to earlier is to reconcile the fact that children's linguistic performance is in some ways so obviously 'child-like' with the

strong arguments for UG as an underlying biological matrix determining the range of possible grammars, including child grammars (Chomsky 1965, 1980, 2001). More specifically, I would like to understand what aspects of children's linguistic behavior (particularly, in the area of the semantics of quantification) can be explained by development and for which ones we should maintain the "as-if-instantaneous" acquisition model.

Abstracting away from the details of various specific proposals, the core generative theory of child language acquisition has largely remained unchanged, maintaining that in order to explain the fact of effortless and universally successful language acquisition on the basis of the deficient primary linguistic data available to the child, we have to posit rich abstract innate linguistic knowledge (UG), which constitutes the initial state of the child's language faculty. Endowed with this knowledge, the child performs the task of building the grammar of her target language by constructing a theory capable of dealing with the input (or in a more modern formulation by setting innately given parameters to their target values).

As explained in Chomsky 1965, this approach belongs to the rationalist school of thought, which holds that knowledge is acquired by humans not simply by means of "peripheral processing mechanisms", as the empiricist school of thought maintains, but by using "innate ideas and principles of various kind that *determine the form of the acquired knowledge in what may be a rather restricted and highly organized way*" [emphasis is mine] (Chomsky 1965). This approach has served as a very productive research strategy. By insisting that the properties of possible natural language grammars are determined by UG, it suggested that points of grammatical variation between natural

languages are not random and potentially infinite in number, but conform to a set of restrictions (i.e. principles, rules, constraints). The latter correspond to the universal grammatical properties (UG), which constitute the initial state of the child's linguistic competence. Consequently, we may be able to discern the universal properties (the content of UG) by doing careful cross-linguistic analysis of adult grammars. Furthermore, the child grammar at any observable stage must be limited by these properties and may not contain arbitrary characteristics not found in any (adult) grammar.

One formulation of this principle became known as the Strong Continuity Hypothesis, according to which "[in] the absence of compelling evidence to the contrary, the child's grammatical rules should be drawn from the same basic rule types ... as the grammatical mechanisms of adults in standard linguistic investigations." (Pinker 1984).

A more recent articulation of this principle is the Strong Uniformity Thesis, which maintains that each attainable state of the language faculty is a further specification of the initial state with parameters fixed. (Chomsky 2001). According to this, 'stages' of linguistic development do not involve a random quantitative or qualitative change, but embody a restrictive process of fixing values for a finite number of choice points left open in the UG within a highly circumscribed search space.

Adopting this approach has been very productive in uncovering a large body of child language data showing evidence of the early knowledge of abstract and sophisticated grammatical principles. Evidence has been presented for children's early knowledge of such abstract principles as Subjacency in English-speaking children (Otsu 1981), c-command and locality as binding conditions for Principle A (Chien and Wexler 1985), structure-dependence (Crain and Nakayama 1986), differentiation of empty

categories in different embedding structures in Japanese (Lust, Wakayama, Mazuka and Snyder 1985), functional categories and clause structure (e.g. Poeppel and Wexler 1993, Deprez and Pierce 1994 among others) to name just a few examples (for a more comprehensive survey see Crain and Thornton 1998, Lust 1999).

One particularly interesting observation has been that it is unusual to find structures randomly different from those found in adult language in either children's spontaneous production or in experimental data. In those rare instances when children do exhibit unusual (for their target language) structures, the error was shown not to be random, but to constitute a grammatical option for some adult-language. Thus, the error of inserting a medial wh-phrase in long-distance questions reported for English–speaking children, is a grammatical option found in certain German dialects and is not confined entirely to child language (Thornton 1990). Moreover, such notorious feature of the early speech of English-speaking children as absence of subjects, coupled with a lack of overt expletives like *it* and *there,* as well as auxiliary *be* and modal auxiliaries, has been argued to constitute a stage at which English-speaking children have the 'null subject parameter' set to the null subject value, a UG-permitted option found in adult Italian- and Spanish-type languages (Hyams 1984).[1] Observations of this kind are used as arguments for the Strong Uniformity assumption.

Along with the strong evidence for UG and the Continuity assumption, there is substantial empirical evidence that certain parts of child language undergo development. Thus, research showed that English-speaking children have difficulty interpreting and

---

[1] Hyams' analysis has been criticized and has undergone some revisions. The major alternatives to the 'parameter setting' accounts have been performance accounts (e.g. Valian 1991, Boster 1997, etc). Lillo-Martin 1991 provides important evidence from ASL

producing passives, particularly those derived from non-actional verbs and those that contain a by-phrase (Maratsos et al. 1985).[2] In response to these findings, the "A-chain maturation" hypothesis was put forward proposing that children initially cannot assign theta roles to arguments moved out of their position because they cannot represent argument–trace chains (Borer & Wexler 1987). Similarly, it was shown that young children have difficulty accessing inverse scope interpretations for doubly quantified sentences, and some researchers then claimed that they go through a developmental stage in which children obligatorily interpret scope-ambiguous sentences in a way isomorphic with their surface syntax (Musolino 2003). According to some of the most radical maturation proposals, the structure of very young children's utterances consists of maximal projections of lexical categories (i.e. akin to the notion of the small clause in the sense of Stowell 1981), either in total or partial absence of functional categories, such as agreement and tense (Guilfoyle & Noonan 1988, Lebeaux 1988, Platzack 1990; Radford 1990). Others argue for the availability of one or two functional projections; e.g. a single category resembling what will become IP, or both AgrP and TP but not CP (Clahsen 1990, Gawlitzek-Maiwald et al. 1992, Meisel & Muller 1992). Under this view, the developing grammar gradually approaches the adult state by adding functional categories to the child's syntactic inventory step by step. Other theories appealing to biological maturation as a major mechanism by which a child turns into an adult linguistically include certain theories of Optional Infinitives, i.e. a lack of consistent inflection in very young children's speech (e.g. Wexler 1995a, 1995b, Rizzi 1994[3]), as well as the maturation of the left periphery (Grinstead 1995).

---

[2] See, however, evidence to the contrary from O'Brien, Grolla and Lillo-Martin 2006

[3] In his more recent work, Rizzi revised his "Truncation Hypothesis" within the Parameter Setting approach.

In the debate between the proponents of Strong Continuity and those of Maturation, the latter argue that Maturation is the simplest, most parsimonious and biologically plausible way to describe child data of the type mentioned above. They argue that since children undergo biological maturation in other areas, it is likely that their 'UG-access' also undergoes maturation (e.g. A-movement, or QR). Critics, however, argue that maturation lacks explanatory force since it merely describes the behavioral facts of acquisition by positing ad hoc rules without explaining "why UG should be so fractionated, why it would appear in parts as it is proposed to do, and why the parts are ordered as they are" (Lust 1999, p. 125). Also, by accepting maturation of various parts of UG, we depart from the foundational hypothesis of UG as the initial state of language acquisition and arrive at a weaker one of UG as the final state of acquisition. Of course, whether or not maturation of UG takes place is an empirical question and cannot be resolved solely by theoretical arguments. However, given the unrestricted theoretical power of the maturation hypothesis, the most fruitful research strategy would require that those explanations that use grammatical mechanisms limited to children and not found in adults be accepted only after no Continuity explanation for them can be found. As put by Pinker (1984):

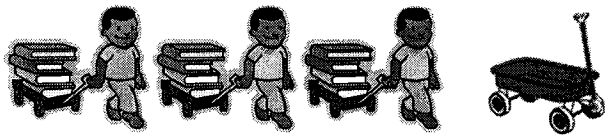
> " There could be an ad hoc child grammar, with no resemblance in form to the adult grammar, plus a mechanism that simply expunges the entire child's grammar at a given stage and simply inserts full-blown adult grammar in its place (to be interpreted, perhaps, as a consequence of neural maturation)... [However,] in the absence of compelling evidence to the contrary, the child's grammatical rules should be drawn from the same basic rule types, and be composed of primitive symbols from the same class, as the grammatical rules attributed to adults... This is not a dogmatic denial that maturation affects language development, just a reminder that there is a burden of proof on proponents of maturational change."

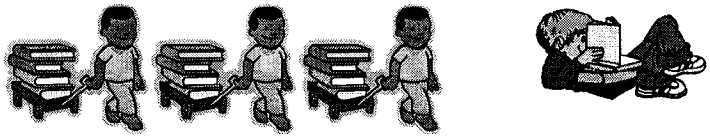
I will follow this approach in my study of one area in which children are reported to be different from adults, namely the semantics of quantification. I will show that the reported differences between adults and children in judging the truth value of universally quantified sentences can be accounted for without positing semantic or syntactic discontinuity, but by positing development outside the computational system, in an area that interacts with the grammatical knowledge to derive meaning of expressions and has been independently shown to undergo maturation, namely Theory of Mind.

1.2. Quantifier-Spreading Errors

It is well attested in a variety of languages that children exhibit non-adult-like behavior with respect to interpretation of universally-quantified sentences like (1) and (2) under certain experimental conditions (an error that has been labeled "quantifier-spreading" or q-spreading, the term I will continue to use for convenience). Thus, it has been found that while children consistently judge (1) true when it is presented in the context of a picture, in which every boy is paired with a wagon and every wagon is paired with a boy, if an extra wagon is added to the picture, as in (3), children often reject (1) as false. Similarly, while they usually accept (2) when the sets of boys and of wagons pulled by them fully overlap, children may reject it in a situation with an extra, non-wagon-pulling boy, as in (4) (Philip 1995).

(1) Every boy is pulling a wagon.
(2) A boy is pulling every wagon.

(3) 

(4) 

As I will review in more detail in chapter 2, one influential theory of q-spreading, which goes back to the pioneering work of Inhelder and Piaget and whose most prominent modern adaptation was developed by Philip 1995 (following Roeper and de Villiers 1993), attributes the errors to a stage in children's development of the semantics of the universal quantification at which in order to verify the truth of such sentences, children have to check whether a symmetry holds between the agent and the theme sets (or the subject- and the predicate sets for Inhelder and Piaget). Thus, according to this approach, in order for sentences like (1) and (2) to be judged true, children check whether all of the boys ride ponies and all of the ponies are ridden by boys (furthermore, for some children the relevant scene must be limited to the fully overlapping sets of boys and ponies and not contain any other participants of any kind). The question I will try to answer is the following:

(5) Are we forced by the available data to accept the view that children go through a developmental stage when their semantics of quantification, particularly that of the universal quantifier, is different than that of adults'? Alternatively, can such non-adult-like behavior be attributed to a weakness outside of grammar, particularly an area of extra-linguistic competence, whose stage-like development is independently motivated and can be straightforwardly accounted for by neurological maturation?

The extra-linguistic area I have alluded to at in (5) is the so-called Theory of Mind or ToM, an autonomous cognitive module responsible for the uniquely human ability to

reason about the content of other minds; e.g. to be able to infer others' desires and beliefs in a given situation fairly accurately (Fodor 1992). The idea I explore in this dissertation is that since semantics of many expressions involves 'mind-reading', children's weakness in the ability to 'read' another person's mind may have direct and observable effects on how well they can interpret expressions relying on this ability. For example, lacking a fully developed ToM would be linguistically significant when having to judge whether a pronominal reference has been made clear to the listener or in making an assumption about the listener's ability to choose the correct interpretation of a context-sensitive expression.

The type of contextual sensitivity I will be interested in has been formally expressed as a contextual parameter, used to account for context dependency of certain expressions, like *everybody* or *only*. A standard proposal is to assume that such expressions contain silent pronoun-like elements (contextual variables) that receive a value from context (Westerståhl 1984, von Fintel 1995, Marti 2003). Consider the following example:

(6) Everybody$_C$ had a great time.

This sentence must be interpreted in relation with a particular context specifying the domain for the quantifier. For example, the speaker may be relating her experiences of the night before, when she and some of her students went out for pizza. By uttering (6), the speaker does not intend to convey the idea that everybody in the whole world had a great time. Rather, the quantification is over a restricted set of people, those who went out for a pizza last night with the speaker. Positing a contextual variable (C) in the lexical entry of the quantified determiner *every* is one way to capture this context dependency. A

common notational device adopted by this approach is a subscript 'C' as in "everybody$_C$".

When using utterances like (6), the speaker makes an assumption that the listener can assign a correct value to the domain restriction variable, e.g. through integrating overt linguistic context or using inference, real world knowledge, etc. Adult ability to make accurate assumptions about what information is available to one's interlocutors in such circumstances is so basic that it is often taken for granted in the literature dealing with acquisition of these types of phenomena. The question that is generally asked is whether children possess the knowledge of semantic and pragmatic rules, e.g. whether they know that the universal quantifier's domain is contextually restricted, and, therefore, a quantified sentence requires context in order to be interpreted. However, I believe it is important to isolate semantic and pragmatic competence from another type of extra-linguistic competence involved here – being able to judge how much the listener can infer from a given discourse situation and whether the situation fulfills the requirements for interpreting a given expression in a certain way. Children are notorious for being weak in making such judgements. In the next section, I will review some pertinent research in children's development of ToM in order to propose what specific linguistic consequences may result from an immature ToM.

1.3. Theory of Mind

Even though there is no general consensus in the field of ToM acquisition as to what mechanism accounts for the child becoming fully adult-like with respect to ToM,

or even what range of abilities must be included in this module, I will adopt the following fundamental assumptions supported by considerable evidence:

(i) full-fledged ToM is unique to humans,
(ii) is genetically encoded,
(iii) has a complex and modular internal architecture, and
(iv) follows a universal progression in the emergence of its various pieces.

The evidence for the innately-specified cognitive module responsible for our psychological reasoning and a developmental time table for its emergence include research showing that certain pieces of cognitive machinery that allow humans understand the minds of other humans are present at birth, e.g. being able to differentiate stimuli with characteristics of a human face and voice from other objects and sounds and showing a preference for the former (Spelke & Cortelyou 1981).[4] Others emerge in the first year of life, such as joint visual attention – pointing, following someone else's pointing, and finally following another person's line of gaze with one's own gaze – which are in place between 8 and 11 months (Scaife and Bruner 1975, Butterworth 1991). In the 2nd and 3rd year of life children develop the ability to report propositional attitudes in their spontaneous speech, first those of desire and later of belief and knowledge (Bartch and Wellman 1995). Another major milestone arrives between 4 and 5, when children become successful on standard False Belief tests (Wimmer and Perner 1983).

[4] This ability along with infant predisposition for being more interested in the human-like stimuli can be viewed as part of ToM because it points to our innate knowledge of the human and non-human distinctions, a chief one of which is human intentionality (due to the fact that humans possess minds), of which infants seem to be aware.

Even though, the bulk of research in this area has concentrated on children's ability to represent false belief, held up to be a litmus test of ToM, other types of developing competence have been identified. Thus, one aspect, which, I believe, is of relevance to the phenomena I am interested in is children's ability to attend to others' access to information and to use it to monitor and update their representation of others' mental states. Research has shown that children have difficulty extrapolating the contents of one's beliefs and knowledge from observing one's access to information. Some researchers found this problem to be so extreme that they claimed that children under 4 years of age, "have no idea of where knowledge and belief come from" (Wimmer, Hogrefe, and Sodian 1988).

This statement, however, is clearly too strong. That children are equipped with certain elements of this knowledge in their first years of life is clear from the "showing" research. Lempers et al. (1977) tested children of 1 to 3 years old on various showing tasks, when the child and another person were seated opposite each other. The child was simply handed either a toy, a picture, or a block with a picture glued to one side, and asked to show it to the other person. Even 1 year-olds were able to hold the toy up to show it. However, until 1;6-2, children oriented it to face themselves. Similarly, until 2, children held the picture horizontally between themselves and the observer, and only at 2;6 for the picture and 3 for the block picture did they orient the object the right way for the observer to be able to see it. Children between 2;6 and 3 were also able to figure out that if a screen was placed between them and the observer, they needed to move it aside for the observer to see the object. This research shows that children become progressively more and more sophisticated in their understanding of people as observers.

They do not however, become fully adult-like in this skill at 3, and their development in this domain continues until at least 8-years-old.[5]

In one experiment investigating children's ability to make correct connections between sources of information and their epistemic effects on others, two children (one of whom was the subject of the experiment and the other played the role of an observer) were seated at a table facing each other. On the table there was placed an opaque non-descript box containing some object, like a pencil or a bar of chocolate. On each trial, one of the two children was shown the contents of the box, while the other couldn't see inside the box. Then, the subject was asked whether she knew what was in the box and whether the other child knew it. They found that 3- year-olds claimed to know even without having seen the contents of the box, giving "yes- responses" not only on those trials when they were shown it, but also on those when it was the other child who was shown it. At the same time, they denied the knowledge of the other child on both the trials when that child did not see what was inside, as well as those in which he was shown what was inside (Wimmer, Hogrefe, & Perner 1988). Furthermore, when the experiment was replicated with an added question of whether the other child had looked in the box, the subjects responded correctly to the look-question, while making errors with the know-question.

Similar results were obtained with older children with respect to inference as a source of one's knowledge. While relying on inference in their own acquisition of knowledge, 4- and 5-year-olds denied that the other person might know via inference.

---

[5] Most likely this type of reasoning continues to develop beyond that age and shows individual variation in adults. This is the type of reasoning that H. Gardner described as Interpersonal Intelligence, one of the multiple intelligences that show variation across people due to a large genetic component.

In this experiment, the subject and the other child examined the contents of a container together seeing that it contained multiple objects of one kind only (e.g. marbles). Then the subject was shown that one item was taken from the container and put in an opaque bag. The other child was not shown the transfer, but was told that one of the things from the container was placed in the bag. The subject then was asked whether the other child knew what was in the bag. Most 4-year-olds responded with the type of response termed "inference neglect" – they denied the other child knowledge through inference, in contrast to their own successful use of inference in another condition (Sodian and Wimmer 1987).

Another demonstration of the difficulty children have with understanding informational access was done by Taylor 1988. In her study, she tested children ages 3- to 8-years-old. Each child was shown a picture of two animals, which were subsequently covered with only a detail showing through the cover. Then the child was asked whether a puppet (who the child was told had not seen the picture before) knew whether the picture contained the animals by looking at the picture with a cover on. The question was repeated with the visible part of the picture of varying size – an empty one, one showing only an edge of a line, one with a small non-descript part of one of the two objects, one with small non-descript parts of both objects, and finally one with a part sufficient to identify one of the objects. Results showed that only the 8-year-old group consistently gave correct answers, denying knowledge in all but the last condition – when an identifiable part of the object was showing. Children in all other age groups (3- , 4-, 5-, and 6-year-olds) over-attributed knowledge in all but the empty conditions.

It's interesting that the two studies described above (Sodian and Wimmer 1987 and Taylor 1988) showed that children err both on the side of under-attribution (as was the case in S&W) and over-attribution (in Taylor) of knowledge to others. I will not try to explain the difference between the two studies here leaving it to future research. The conclusion they lead me to is only that tasks when children have to figure out what is known to others are hard for them, causing them to use some faulty strategy, which may vary across tasks.

The majority of researchers in this field focused on children's understanding of epistemic effects of visual information (or inference about visual information), as the studies reviewed above. However, further research is needed to investigate how much grasp children have on the epistemic effects of linguistic context on another person. In the light of the previous studies, however, it is plausible to expect significant deficits in being able to make correct assumptions about what another person would know after being exposed to a certain context. As I have suggested, this deficit has interesting implications for research in children's semantic competence since many aspects of interpretation rely on the interlocutors' ability to "read" each other's mind, i.e. to understand what information is salient, relevant, and/or sufficient for one's listeners to be able to interpret one's contributions to discourse.

1.4. Semantic Effects of ToM: Binding Theory and Interpretation

One of the most widely studied areas of acquisition in which developing ToM has been implicated was Binding Theory, particularly the errors children have been reported to make with respect to Principle B (see Avrutin and Wexler 1992 for a review of the

literature on the topic). It had been documented that children, unlike adults frequently accept as true the following sentence in the following context (Wexler and Chien 1985):

(7) context: a picture showing Mama Bear, who is scratching herself, and Goldilocks.
(8) This is Mama Bear, and this is Goldilocks. Mama Bear scratched her.

One obvious approach would be to suggest that children go through a developmental stage in which they do not obey Principle B (in the sense of Chomsky 1981). Since, however, children were shown to be highly successful with Principle A (Wexler and Chien 1985), the authors rejected the maturation of Principle B account, and instead suggested a pragmatic account, according to which children's errors would fall into the category of what I classify as ToM-related errors. In this theory, the children's odd interpretation of (8) is not the one that violates Principle B, in which *her* and *Mama Bear* are co-indexed (as shown in 9), but the one in which they are not co-indexed but nevertheless co-refer, as in (10):

(9) #Mama Bear$_i$ scratched her$_i$.
(10) Mama Bear$_i$ scratched her$_j$.

W&C argued that for the two counter-indexed expressions to co-refer is a grammatical option, which, however, requires deictic interpretation of the pronoun. Adults know that what W&C refer to as "deictic grounding" can only be achieved if there is a means of making the pronominal reference clear to the listener (e.g. with the speaker pointing to Mama Bear when uttering 10). Hence, they rule out (8) in the situation (7) since no such means is available, as they rule out the co-indexed interpretation in (9) due to Principle B and hence reject the sentence. Children, on the other hand, while also being able to rule out the representation violating Principle B, do not detect the unavailability of (10) because they believe that "listeners can infer more from a discourse situation than in fact

they can infer, that is more than is adequately grounded." (Wexler 1999). According to this explanation, the children who commit the apparent Principle B violations, err in attributing greater inferential power to the listener than is warranted in the situation for reasons independent from their syntactic or semantic competence. This type of error is what I suggest is a manifestation of the developing ToM, responsible for a wide range of interpretive errors that at first glance may appear to be semantic in nature.

As evidence that the error is not a violation of Principe B, C&W showed that the same children who accept (8) in context (7), correctly reject (11), while accepting (12):

(11) *Every bear$_i$ scratched her$_i$.
(12) Every bear$_i$ scratched her$_j$

Here, while (11) is ruled out by Principle B, which children seemed to obey, (12) is permitted but cannot have a bound interpretation irrespective of the pragmatics of the situation. When the deictic option for co-reference is ruled out by the grammar, children do not commit 'principle B' violations.

1.5. Another candidate: Scalar Implicatures

Another phenomenon that relies on the 'mind-reading' abilities of the interlocutors is scalar implicatures (SI), first discussed in Horn (1972). SI arise with certain expressions, interpreting which requires the listener to make assumptions about the interlocutor's communicative intentions. Thus, the smoothness in the mini-discourse in (13) relies on speaker B's knowledge that in addition to the literal meaning of A's utterance, which truth-conditionally would be fully compatible with the situation in which the property in question is true of all journalists, it carries an implicature denying

that *all* journalists care more about getting invited to the Christmas party at the White House than they care about the truth.

(13) A. Some journalists care more about getting invited to the Christmas party at the White House than they care about the truth.
B. I am glad not all of them do.

According to a standard analysis, figuring out the implicated meaning involves knowing that semantics of A's utterance includes a scale *<some, all>*, where the second member of the scale is informationally stronger than the first, as well as knowing the pragmatic principle requiring that the most informative member of the scale believed by the speaker to be true be used. Then, since the speaker did not use the more informative *all,* one can infer that the speaker knows it to be false that *all* journalists care more about being invited to the Christmas party at the White House than they care about the truth.

This implication, however, is defeasible. Thus, if the context entails that the speaker does not have enough evidence, e.g. has only a partial access to information, or simply chooses to focus on a subset remaining agnostic about the rest, the implication is defeated and *some* becomes compatible with the informationally stronger *all*, as in (14). B's answer is given on the basis of his observation of a subset of journalists (namely those who asked questions at the press conference) and may possibly apply to all (relevant) ones, which allows for A's reply to follow smoothly, even though his remark is more general than B's, and would contradict the implicature.

(14) A: Did you listen to the White House Press conference yesterday? No one asked any tough questions.

B: Yes, clearly some journalists care more about getting invited back to the White House than they care about the truth.

A: I agree, the only thing they all worry about is loosing access to the White House officials.

A crucial factor that plays a role in defeasibility of SIs is the communicative intention of the speaker to convey the implicature. Notice that in (14), we can't detect such an intention. This makes it possible to assign the sentence a reading in which the communicatively weaker item doesn't rule out a stronger one. If the speaker has an intention to communicate the SI, and if the context makes the intention clear, it becomes difficult to suspend it:

(15) A: How was the conference?
B: Some talks were interesting.
A: #I am glad you liked it so much.

Since the question in (15) asks about speaker B's opinion about the entire conference, B's response is intended to be about the entire conference. Since the speaker deliberately substituted a weaker expression for a stronger one, this allows the listener to read the substitution as a negation of the proposition applying to all talks at the conference; hence, the oddness of A's response.

It is clear that for interpreting SI, children would need to be capable of a highly sophisticated reasoning about communicative intentions and access to evidence by others, the skill they may be expected to be weak at, given the evidence discussed in the previous section. Furthermore, it is reasonable to suggest that this weakness is independent from their semantic knowledge of scales. Even if children are not different from adults with respect to knowing the mechanism of computing scalar implicatures, their immature ToM may lead them to non-adult-like performance on tests measuring their ability to compute implicatures, particularly when the experimental design makes heavy demands on their ability to read speaker intentions. This hypothesis is consistent with the experimental data with respect to children's knowledge of scalar implicatures.

Some studies have reported that children have difficulties computing implicatures (e.g. Chierchia, Crain, Guasti, Gualmini, and Meroni 2001; Noveck 2001; Foppolo and Guasti 2004). However, other research showed that with improved experimental design 4-6-year-old children can do so successfully (Papafragou and Tantalou 2004). Explaining the dramatically improved success rate in computing scalar implicatures, P&T point out that in those studies in which children failed to show sensitivity to SI, the sentences containing a weaker scalar items in contexts where the stronger one was true, were presented by the "ignorant" puppet, who failed to notice that his statement carried the potential for conveying SI (Papafragou and Musolino 2003). In other such studies, no context was provided (Noveck 2001), which left it up to the children to reconstruct possible scenarios. Both designs leave open the option discussed above when a weaker scalar item can be interpreted as compatible with its stronger counterpart. On the other hand, in the P&T 2004 study, the experiment was carefully designed to address the pragmatics of SI.

In their experiment, animals were told to perform various tasks, such as coloring a number of stars, eating a number of sandwiches, or wrapping a number of gifts. The animal then went into a dollhouse to perform the task out of sight. After a while, he came out and was asked by the experimenter whether the task was complete:

(16) Experimenter: Did you color the stars?
Elephant: I colored some.

After hearing the animal's response, the child was to reward him if she thought the job was done. In this task, the speaker's (elephant's) communicative intention and access to evidence was controlled to insure that the targeted response was dependent on a successful computation of the SI. The elephant was to report whether he had succeeded

coloring all of the stars. Hence, reporting that he colored some of the stars was perceived as a deliberate response and resulted in a dramatic improvement in children's success: they responded consistently with a computed SI 70-90% of the time across different experimental conditions.

P&T conclude that their results provide compelling evidence for children's early ability to compute implicated aspects of speaker's meaning. Given that the context sets up clear expectations about how informative the target sentence must be, they interpret the failure of the target sentence to meet these expectations as conveying an SI.

This discussion serves to illustrate a problem in semantic acquisition research, when experiments designed to test children's competence of the semantics of a certain expression fail to insure that it is presented in such a way that it does not only avoid pragmatic infelicity, but sets up unambiguous pragmatic conditions for the targeted expression. The necessity of such careful pragmatic scaffolding can be easily overlooked because typically adults easily overcome pragmatic "gaps" by being more skilled test-takers and hence being able to read experimenter's intentions much better than children. Thus, in Papafragou and Musolino 2003, adults rejected under-informative statements over 90% of the time, while 5-year-olds only 12.5% on the some/all condition. The difference in being able to compute SI between children and adults in that experiment may likely to be due to their respective abilities dealing with pragmatically ambiguous situations. In the experiment, the speaker is a 'silly puppet', who utters under-informative sentences presumably out of ignorance. His ignorance can be manifested in conveying a false SI or a false sentence, both of which would have to be rejected. However, this expectation of the speaker's ignorance can be counterproductive in an

experiment targeting conversational implicature because, as discussed above, the implicature is calculated precisely on the basis of the speaker's perceived intentional use of a weaker term. Since the under-informative sentences were truth-conditionally correct, it is not surprising that children accepted it, particularly when contrasted with the false fillers.

2. Hypothesis

I have proposed that weaknesses in Theory of Mind have linguistic consequences, i.e. they may cause children to misinterpret those linguistic expressions whose interpretation involves accessing the state of mind of one's interlocutor. In addition to Binding and SIs discussed in the previous section, other expressions also require this ability, particularly expressions containing contextual variables, such as quantifiers.

The hypothesis I wish to put forward is that the typical quantificational errors should be treated as a result of miscommunication due to children's developing ToM and not a manifestation of non-adult-like semantics of universal quantification. In order to pursue this theory, I will evaluate the major earlier findings in this area and present the results of the experiments I have conducted, which I believe shed some new light on the nature of children's errors. Taken together with the previous core findings I will review and analyze, they should help us address the question that I posed at the beginning of this chapter, namely whether or not the well-attested errors with respect to quantification constitute the sort of compelling evidence for abandoning the Continuity assumption that were asked for by those who argued for Continuity as an important constraint on the theory of language acquisition.

In the debate about the proper treatment of children's grammar of universal quantification, there are those who claim that children go through certain pre-adult semantic stages before reaching full semantic competence, and those who argue for early full semantic competence. In this debate, both sides have collected a large body of empirical evidence in support of their respective view; hence leaving the puzzle largely unresolved. The thrust of criticism leveled by either side against the conflicting findings has been to attribute them to a flawed experimental procedure, which either artificially creates errors as a result of pragmatic infelicity of how the test is administered (Crain et al. 1996) or 'hides' those parts of the context that would normally cause the child to give a non-adult-like response by making them less salient or in some way pragmatically irrelevant (Drozd and van Loosbroek 1998, Gordon 1996).

In looking for an explanation for why children make the errors in question, I will examine the core experimental findings that have been reported in this area. While some previous accounts tried to explain away the findings that did not fit well with their theory, I will try to unify the core data by finding a common explanation. I will adopt the Continuity Assumption, but not as a dogma to be strictly followed even at the expense of ignoring important empirical evidence, but as a research strategy, which requires one first to look for ways to explain children's behavior using known adult grammatical mechanisms, to be revised only if such mechanisms are demonstrated to be empirically inadequate.

After carefully considering which experimental conditions engender errors and which ones minimize them, I will formulate a hypothesis, looking at the role of the ambiguity of the sentences like (1) and (2) with respect to specific/non-specific indefinite

interpretation and the pragmatics of the typical experiments in which children produce the errors in question. I will suggest that such errors arise when the child has to interpret expressions that involve contextual parameters in a situation when sufficient context has not been provided.

According to my hypothesis, the key to explaining children's behavior with respect to sentences like (1) and (2) in the contexts illustrated in (3) and (4), is not in their knowledge of the semantics of the universal quantifier, but in their interpretation of the indefinite NPs, namely in how they use visual context to restrict the domain of indefinites. The typical quantified sentence used in the tests eliciting quantification errors has an ambiguity largely ignored by previous research. This ambiguity is due to the presence of two scope-bearing elements – the universal quantifier *every* and an indefinite NP.[6] In addition to the syntactic interaction, there exist multiple options of how the domain of the indefinite may be restricted.[7] To illustrate these ambiguities, let's consider sentence in (17). It is ambiguous between two readings, one with the truth conditions as in (18), in which the universal takes wide scope, and the other with those as in (19), with the indefinite taking wide scope:

---

[6] I will assume an approach to indefinites, originating from Russell (1919), that treats indefinites as introducing an existential quantifier; i.e. a sentence "a dog barked" has the meaning 'there exists some relevant individual x, such as x is a dog and x barked.' This view is in opposition to the view that treats some indefinites as referential (referring to an individual rather than making an existential statement). Under the Russellian view, all indefinites are scope-taking quantificational elements, that interact with other scope-taking elements, such as other quantifiers, negation, etc.

[7] The idea that quantifiers come with unpronounced domain restriction variables ranging over properties of individuals was first introduced by Westerståhl (1984), and further developed in von Fintel 1994, Schwartzschild 2001, Martí 2003 among others. The role of the domain restrictions can be illustrated with the following example. The truth-value of a sentence like 'everyone is unhappy' is always evaluated against some contextually determined domain, which includes a set of relevant individuals only all of whom must be unhappy in order for the sentence to be true. For example, if this sentence is uttered in the context of a department faculty meeting, it could refer to everyone in the department, every faculty member in the department, or every person present at the meeting, etc., depending on the specific context. Thus, the existence of a happy student would not falsify the sentence if the domain is narrowed to include only the department's faculty.

(17) Every boy is riding a pony.
(18) For every x, if x is a boy, then there is a y in C, y is a pony, and x rides y.
(19) There is a pony y in C, such that for every x, x is a boy, x rides y.

If (17) is interpreted in such a way that the domain restriction C for the indefinite *a pony* includes multiple individuals y, the narrow scope indefinite interpretation in (18) would allow a co-variation between members of the sets of X's and Y's (i.e. a situation when every boy is riding a different pony, a reading I will be referring to as non-specific). If, on the other hand, the domain C includes only one unique pony y, the sentence would receives an interpretation requiring that every boy rides that unique pony (henceforth, I will refer to this reading as specific or singleton indefinite reading) regardless of the syntactic scope relations in the sentence, and both (18) and (19) would have identical truth conditions. [8] We can see this ambiguity, if we add a continuation as in (21) and (22) to (20):

(20) The boys went to the stable full of ponies. Every boy rode a pony.
(21) The tireless pony took each boy around the meadow and brought him back to the stable.
(22) Each pony had ribbons and a saddle blanket of a different color.

Even though the scenario contains multiple ponies, in (21) the domain of the indefinite is implicitly narrowed to include one relevant pony, while other ponies mentioned in the scenario are excluded. In (22), the domain of the indefinite includes multiple ponies. In a natural discourse, the interlocutor would use contextual cues to understand when the domain has been narrowed, and would avoid a "garden-path" effect at the point of interpreting the subject in (21) or (22).

---

[8] I will assume that the presuppositional semantics for 'every'.

The crucial point here is that the sentence is potentially ambiguous and requires context in order to be disambiguated. This fact have been overlooked in the Q-spreading literature probably because for English-speaking adults, sentences with universal subjects and indefinite objects (and no negation) by default receive the non-specific interpretation in the absence of a context highlighting one relevant individual, and the singleton interpretation is a marked option (Kurtzman & MacDonald 1993). The ambiguous sentence presented without linguistics context (or with insufficient linguistic context), where the participants are introduced as a minimally informative visual scenario (which shows the participants, but fails to show which ones are the relevant ones and which ones are there only as distracters), presents a challenge to the child, who has not yet become a sophisticated taker of psycholinguistic tests.

I propose that making this decision is where his weakness lies. The child as young as 3- and 4-years old has a sophisticated knowledge of semantic and pragmatic rules, which includes not only the correct meaning of the universal and existential quantifiers and how they interact with respect to their relative scope, but also of the role of context in delineating appropriately narrowed domain restrictions. The difficulty, however, arises when she has to make decisions about fixing contextual parameters for context sensitive expressions in conditions of contextual paucity. When faced with an ambiguous sentence under such conditions, the child is likely to choose an interpretation allowed by the grammar but unintended by the experimenter. This interpretation may be so pragmatically odd to an adult, that it would be labeled 'incorrect' or 'semantically anomalous', even though such responses are not semantically anomalous and would be fully grammatical in an appropriate context. This approach, I believe, explains the patterns and the relative rate of the observed errors across experimental conditions.

This explanation makes some predictions, some of which I will test in the experiments presented below. Thus, it predicts that asymmetric contexts and the quantification errors can be dissociated in a number of ways:

- Given sufficient context, preschool children should be able to give correct responses even in situations of contextual asymmetry.
- Since the element responsible for the inappropriate responses is not the universal quantifier, as has been maintained previously, but the indefinite NP, we should be able to see the errors caused by selecting the specific indefinite reading when the non-specific one is intended irrespective of the condition of symmetry. In particularly, such errors are expected in constructions and contexts that have been shown to cause the specificity bias, e.g. sentences where the indefinite NP occupies the subject position and the universal the object position (Kurtzman and MacDonald 1998).
- Likewise, we can expect to find similar non-adult-like responses with quantifiers other than the universal under the experimental conditions that typically give rise to such responses, namely in conditions of asymmetric picture scenarios.

In chapter 2, I will discuss the core findings reported previously and review major earlier theories that tried to account for these findings. In chapter 3, I will present my proposal in detail and discuss how the semantics of indefinites can be responsible for both classic q-spreading errors and children's non-adult-like scope preferences reported in the literature. In chapter 4, I will present the results of three of the experiments I have conducted testing some of the predictions made by my hypothesis. In chapter 5, I will offer a hypothesis regarding the remaining type of errors, namely so-called 'perfectionist errors'.

Chapter 2. Previous Accounts

2.1 Symmetry in children's responses

Since the work of Inhelder and Piaget (1964), it has been known that until the age of 7–8 children's responses are frequently different from those of adults when judging the truth-value of sentences with a universal quantifier, such as English 'every' or 'all'. The core finding is that while children seem to know the meaning of the universal quantifier, and would correctly reject a sentence like 'every circle is red' if a picture shown to them contains a number of red and a number of blue circles, under a closer scrutiny their universal quantifier appears to differ in its logical properties from that of adults'.

The type of sentence that has been studied the most is like that in (0):

(0) Every boy is riding a pony.

The main puzzle concerns children's responses in so-called minimal contexts, where the context consists of a visual array containing a set of boys riding ponies with one or two extra individuals of some type (an extra boy, an extra pony, a girl with or without a pony or a girl riding another type of animal). It has been demonstrated that in their judgments, children often react to the presence of an extra individual in peculiar ways. Thus, children have been shown to be able to judge correctly as true the sentence in the context of a picture containing two balanced sets: that of boys riding ponies and ponies being ridden by boys.[1] At the same time, children of the same ages have been documented to produce so-called 'symmetry error' responses giving a no-answer in situations when the picture contains an extra pony or an extra girl, and 'under-exhaustive'

[1] Some recent results from Altreuter and de Villiers 2005, however, contain a number of children who incorrectly rejected sentences in a context of two exhaustively paired sets. Because of the oddity of such a response, no linguistic explanation has been put forward.

errors - a yes-response if the picture contains an extra boy.[2] These odd responses have been used to argue that children interpret universal quantifiers in a way different from adults.

2.1.1. Symmetry requirement as a cognitive property

The research that first brought to light children's asymmetry triggered errors was done by Inhelder & Piage (1964). They discovered that when Swiss French-speaking children were presented with an array of blue circles and blue and red squares and asked "Are all the circles blue?", they often responded 'no' and as an explanation pointed to the blue squares. Inhelder and Piaget called this response symmetrical, and concluded that for young children (Piaget's stage II), sentences with the universal quantifier require symmetry between the set denoted by the predicate and the set denoted by the common noun of the subject. As Inhelder & Piaget put it, "It ... looks very much as if the true explanation is that at stage II children extend the quantifier 'all' to the logical predicate of the sentence as well as to its logical subject" (Inhelder & Piaget, 1964, pp. 70-71). I&P attributed the errors to immaturity of logical reasoning, namely the child's inability to construct hierarchical systems or operate in terms of a schema of class inclusion. Thus, they claimed that children at this stage, when presented with a question: 'Are all Xs Ys?', where Xs are individuals and Ys are properties, construct two non-graphic collections (as opposed to the more primitive graphic collections stage I children operate with), one of Xs and one of Ys. According to I&P, the child, unable to cope with the problem of the quantification of the predicate, is limited to ascertaining whether or not the collection of Xs

[2] Another type of non-adult-like no-response has been elicited when a picture contains an extra non-boy (e.g. a girl) or a non-boy pony rider (a girl on a pony) and was first discussed by Philip 1995, who termed this type of error perfectionist. Later, a similar type of response was discovered by Roeper et. al (2004), i.e. no-responses given if there is an extra pair of unmentioned objects involved in the same activity as the mentioned individuals (an elephant-riding girl). Roeper et al. called this error type "bunny-spreading" because their example case involved a picture with extra bunnies. I will include 'bunny-spreading' errors into the class of perfectionist errors, which I will discuss in chapter 5.

coincides with that of Ys. This amounts to reducing the question 'Are all the Xs Ys?' to 'Are all the Xs all the Ys?' instead of the correct 'Are all the Xs some of Ys?' Thus, according to this view, the source of non-adult-like responses in children of this developmental stage is their replacement of class inclusion with equivalence.

Following this explanation, this type of error became known in the literature as the symmetrical interpretation error. Views attributing symmetrical interpretation to children amount to a claim that there is a fundamental difference between the adult and the child semantics of the universal quantifier: adults require the truth conditions in (1) (simplified for the ease of exposition), while children require (2)[3]:

(1) $||\text{every}||(A, B) = 1$ iff $A \subseteq B$
(2) $||\text{every}||(A, B) = 1$ iff $A = B$

I&P's theory offered an elegant explanation of the child data by positing a developmental path from the child to the adult interpretation as a transition from being able to handle a logically simpler operation to a more complex one. However, further research revealed new conditions that give rise to interpretive errors in children, which were not consistent with their explanation and required a new account. Much subsequent research took I & P's hypothesis about children requiring symmetry for universally quantified sentences as the basic descriptive generalization to be given a new theoretical explanation. However, the explanation was now sought not in children's general cognitive capacity, but in their grammar.

---

[3] The Inhelder and Piaget study involved the French equivalent of the English 'all'. Most current research with English-speaking children centers on 'every', which is why I use 'every' in the formulas in (1)-(2). Since I&P did not discuss lexical differences between 'all' and 'every' and spoke generally about children's logical abilities with respect to universal quantification, I will assume that the formula they used for 'all' would also apply to 'every', as given in (1)-(2).

2.1.2. Symmetry requirement as a property of non-adult grammar

Linguistic accounts that followed the influential work of I&P, who had adopted the notion of symmetry as a general cognitive principle, attempted to pose a grammatical mechanism that would lead to the symmetrical interpretation. As subsequent research showed, Inhelder & Piaget's formulation of the children's semantics for universally quantified sentences had to be modified to accommodate new data. Thus, after expanding this line of research to sentences with two-place predicates, it was found that the lack of symmetry between the extension of the subject's common noun (e.g. a set of boys) and that of the predicate (a set of pony riders) is not what triggers the highest rate of symmetrical errors for most children, as would be predicted by (2). Rather, it is triggered by the lack of symmetry between the extension of the subject's common noun (a set of boys) and that of the object's common noun (a set of ponies) does. One major study, Philip 1995, reports that for his so-called 'symmetry' group, the high error rate of 56% was obtained on the condition in which the scenario paired with the sentence 'every boy is riding a pony' contained three ponies ridden by boys and one rider-less pony. On the other hand, when the scenario contained three pony-riding boys and a pony-riding girl, this group gave adult-like yes-responses close to 100% of the time, showing that they do not require symmetry between boys and pony-riders, as would be predicted by I & P's account.[4] This type of data could not be accounted for with semantics given in (2); hence, new semantics had been proposed.

There have been various proposals regarding child's symmetrical responses. One major school of thought attributed the errors to the child's non-adult-like grammatical analysis of the universally quantified sentences. Thus, Bucci (1978) suggested that children cannot restrict the

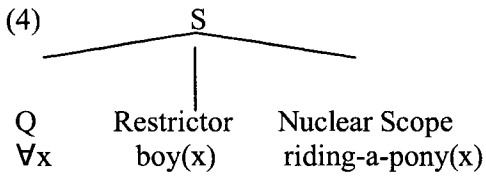
[4] Similarly, Philip's 'perfectionist' children rejected this sentence at a high rate of 85% in the context where an extra individual was a rider-less pony. The same group of children exhibited a much lower rejection rate in the context of an extra pony-riding girl at the rate of 58%.

quantifier to its domain and allow the domain to spread to other parts of the sentence. Roeper and Matthei (1974) claimed that the errors arise because quantifiers *all*, *some*, and *every* are treated by children as adverbs and that children allowed quantifiers to spread from their syntactic position as a determiner to two adverbial positions in the sentence (see also Roeper and de Villiers 1993, Roeper et al 2004 for extensions).

The adverbial view was refined and elaborated by Philip (1995) as the Event Quantification account. He proposed that the universal quantifier for children "alone occupies a position of sentential scope, rather than the whole quantifier phrase", and hence is an adverb of quantification rather than a quantificational determiner (Philip, 1995, p.3).[5] According to this idea, for children, the universal quantifier does not quantify over individuals, as it does for the adults, but quantifies over events. The events forming the restrictor are the sub-events of the contextually relevant event that meet a particular restriction, namely that an individual of either the agent-type or the theme-type is a participant in the sub-event. Thus, a sentence like that in (3), which for adults has a semantic structure as given in (4), for the child, has the semantic structure as represented in (5) and can be paraphrased as in (6):

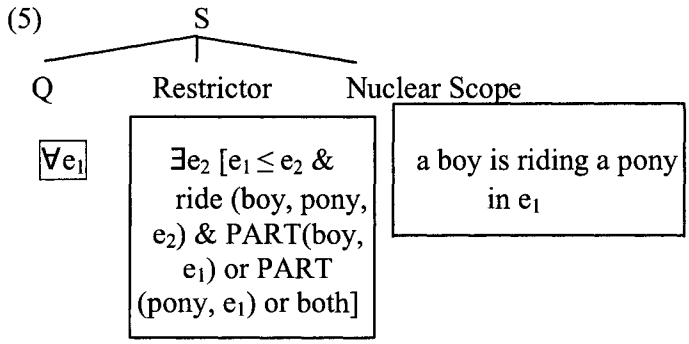
(3) Every boy is riding a pony.

(4)



[5] For convenience, I will refer to the theories seeking a grammatical explanation to the errors in children's performance as the "grammatical symmetry" accounts.

(5)



(6) 'For every event e_1, such that e_1 involves a boy or a pony (or both) as a participant, and e_1 stands in the part-whole relation to an event e_2, in which a boy is riding a pony: e_1 is an event of a boy riding a pony.'

This analysis covers the cases when the child rejects a sentence with either an extra boy or an extra pony. It does not cover another group of children, those who did not only make the former error, but also incorrectly rejected the sentence if the picture contained any unmentioned extra individual, a group termed by Philip 'perfectionist children'.[6] In Philip's theory, to cover this type of error, an additional component had to be added to the restrictor. In addition to stating that participants in the sub-events are boy, or pony, or both, another type of participant had to be added, namely 'any perceived individual'. This addition had the desired result by requiring every sub-event of the event under consideration that involves any participant to satisfy the truth conditions of the nuclear scope. In other words, for the perfectionist child, *every boy is riding a pony* is true if and only if for the event under consideration every sub-event that involves any participant is a sub-event in which a boy rides a pony.

This explanation has been adopted by Roeper et al (2004), who elaborated on the Event Quantification theory and earlier adverbial theories by proposing a syntactic analysis of children's universal quantifier as developing from being an adverb to Floating Quantifier (FQ)

[6] In their later work, Roeper et al. (2004) investigated this type of errors and gave it an informal name 'bunny-spreading'. The name comes from the no-responses elicited when a sentence 'every dog is eating a bone' was paired with a picture containing a number of dogs eating bones plus a bunny eating a carrot. Children accompanied this response with pointing to the bunny and saying, 'no, not the bunny".

before reaching the adult satge when it becomes a determiner. In the acquisition path that they propose, the child moves from the 'bunny-spreading' stage to what they call "classic spreading" stage, to the adult stage, with the final transition occurring in some children as late as the age of 12. According to their proposal, at the first stage, the child does not project a DP above the NP, and interprets subject quantifiers as being adverbial elements in Focus Phrase (FocP), a projection dominating CP, which has been argued for Hungarian quantification (Brody, 1990) and extended to acquisition by Kang (1999). Children at this stage are claimed to interpret *every* as though it were *always.* It takes scope over the entire sentence, quantifying over events in a manner similar to Philip's analysis. Consequently, the presence of any individual other than a pony-riding boy (and conversely a pony being ridden by a boy) would falsify the sentence for the child at this stage of development.

In the second stage, they propose the child syntactic representation becomes more refined, with *every* as well as *each* behaving as Floating Quantifiers, both with the feature [+distributive]. At this stage, the initial quantifier is interpreted raised to Spec-FocP to check its [+dist] feature and transfer its [+dist/+universal] features to the FQ position via co-indexing. At this stage *every* no longer quantifies over events, but over individuals; hence children no longer produce 'bunny-spreading', but still allow classic spreading.

To account for the adult stage, when *every* is reanalyzed as rooted in the DP and no longer acts as a FQ, they integrate Beghelli and Stowell (1997) with the Kang/Brody analysis. They assume movement to a Spec-FocP position dominating CP (following Kang/Brody), but instead of considering the movement as a general operation on quantifiers, they follow Beghelli and Stowell in assuming that the movement is driven by the feature [+dist] found on certain quantifiers but not others. They propose that at this final stage, children learn that *every* is a

mixed quantifier, which is sometimes interpreted as distributive and sometimes as collective. They also reanalyze *every* as a quantifier lacking the feature [+dist], and as a consequence remaining in situ, while its distributive counterpart *each* raises to Spec- FocP. While it remains in situ, the NP containing *every* will sometimes receive a collective interpretation and other times a distributive interpretation, depending on the properties of the predicate that selects it.

Even though accounts of this class capture the symmetrical responses, there are several strong reasons to be cautious about adopting this view. One may argue that 'grammatical symmetry' theories underestimate child competence since along with the symmetrical response data persisting until as late as 12-years-old (Roeper et al 2004), researchers have obtained adult-like responses in children as young as 3 (Crain et al. 1996, Philip 1995). In addition, contrary to the predictions in Roeper et al. 2004, children have been demonstrated to have non-distributive readings of universally quantified sentences (Brooks and Brain 1996, Rakhlin 2004). Another reason for skepticism with respect to a grammatical account is that the available experimental data clearly shows that the symmetrical errors are limited for the most part to certain types of experimental conditions. The tasks resulting in high rates of errors are those in which the test sentences are given without or with a very limited linguistic context. Instead, children are shown a visual array of individuals and asked whether the truth or falsity of the test sentence holds of these individuals. In contrast, when the test involves a rich linguistic context, the error rate is dramatically reduced (Crain et al. 1996, Drozd and van Loosbroek 2001, Rakhlin 2004).[7] Finally, the adverbial quantification theories do not account for the 'under-exhaustive' responses - erroneous yes-answers (e.g. accepting as true a sentence such as 'every boy is riding a pony' when one of the boys in question is not riding a pony). These reasons motivated a new line of explanation, namely one attributing to children a full grammatical knowledge of universal

[7] I will discuss the effect of experimental conditions on the rate of errors in more detail in section 2.2.

quantification, but a deficit outside of syntax/semantics.

2.1.3 Weak Quantification Theories

One alternative theory attributes the errors to children's treating the strong quantifier *every* as weak (Drozd and van Loosbroek 1999, Geurts 2004). The basic distinction between strong and weak quantifiers goes back to Milsark (1974), who noticed that while one group of quantifiers, namely numerical, *no*, *few*, *many*, *some*, etc. could occur in existential constructions in English, others, like *every* and *most* couldn't. The former group of quantifiers became known as 'weak', while the latter as 'strong'. As Geurts explains, the key semantic difference between the two classes is that strong quantifiers are inherently relational, while weak ones are not. Thus, expressions of the form 'Most X are Y' always require looking at the entire set X and ascertaining its relation with set of Ys. In this case, it is verifying whether a majority of individuals in X are members of Y. On the other hand, 'Some X are Y', although it can receive an analogous interpretation, does not require it and when ascertaining its truth value, one only needs to consider the intersection between sets X and Y. Thus, the meaning of this expression may be construed as "there exist individuals that have both the property X as well as property Y." Since, in the case of weak quantifiers any members of sets X and Y outside the intersection are irrelevant, while for a strong quantifier one needs to inspect the entire set X and its relation with set Y, interpreting the latter is more complex. Geurts cites some psycholinguistic evidence supporting this observation (Just 1974 and Meyer 1970). Weak Quantification Theories exploit this difference in complexity in processing weak vs. strong quantifiers claiming that the errors children make are rooted in the weak/strong quantifier distinction.

The basic claim made by both D&L and Geurts is that although children's syntax and semantics of *every* is not different from adults in every other respect, unlike adults children treat the universal quantifier as if it is weak. The consequences of this difference are different in the two accounts.

According to D&L, what goes wrong in interpreting a weakly construed universally quantified sentence, is that it gives rise to a different logical inference pattern than is expected in the case of a strong quantifier. Thus, in the case of a weak quantifier, "Q X are Y" is logically equivalent to "Q X that are Y are Y". For a child who is evaluating the truth-value of (3) and mis-analyzing *every* as weak, it would mean "every boy who is riding a pony is riding a pony". This gives rise to non-adult 'yes' responses in contexts of a picture with a non-pony-riding boy, which may now be excluded from the evaluation. On the other hand, the non-adult-like 'no' responses in context with an extra unridden pony arise because of the context-sensitivity of weak quantifiers, namely *many* and *few*, analogous to which children are claimed to treat the strong quantifier *every*. Interpretation of such weak quantifiers involves comparing the number of individuals in the restrictor set to some expected number. To illustrate this, D&L rely on Westerståhl's example in (8):

(8) Many Scandinavians have won the Nobel prize in literature.

As Westerståhl (1982) observed, under the most salient interpretation, (8) means that the number of Scandinavian Nobel prizewinners is larger than one would expect, statistically speaking. Under this reading, the number of such individuals is not necessarily large in proportion to the total number of Scandinavians, but large in proportion to some average number of Nobel prize winners from a single region. Then, if children treat *every* as weak, they would apply the same logic interpreting (3) and would compare the number of pony-riding boys to the expected

number. If they see an additional pony, they expect that it is there because a boy should be riding it, and since that boy is absent, the sentence is judged false.

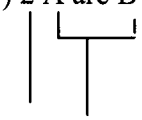
Geurts agrees with the basic claim of weak construal as the cause of quantification errors, but disputes the mechanism proposed by D&L. According to Geurts, the problem lies in children's "malfunctioning mapping from syntactic structure to semantic representation. If this goes off the rails, grammar leaves the domain of the quantifier underdetermined, as compared to adult construals, leaving proportionally more room for pragmatic inferences to determine the eventual outcome of the interpretation process." (Geurts 2004 p. 203). To illustrate how this works, Geurts discusses the following examples (slightly modified):

(9) Two llamas ate an apple.
(10) All llamas ate an apple.

The intended interpretation of (9) is intersectional: there are Q-number of individuals *x* such that *x* has both properties – *x* is a llama and *x* ate an apple. In other words, it would have a semantic form of $<Q>\varphi$. The quantifier binds the first variable in φ; any remaining variables have existential force by default. On the other hand, the meaning of a sentence with a strong quantifier as in (10) is relational: for every individual *x* with the property of being a llama, *x* is included in the set of individuals with the property of having eaten an apple. This sentence has the semantic form $\varphi<Q>\psi$. The quantifier binds the first variable in φ; the rest have existential force by default.

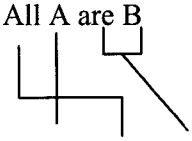
The crucial step in this account relies on a theoretical assumption made by Geurts about an independent level, in which syntactic and semantic forms are mapped onto each other. Thus, the syntactic form of (9) is mapped onto the semantic representation with is as shown in (11), while for sentence in (10), the mapping between syntax and semantics is more complex and is something along the lines of the one shown in (12).

(11) 2 A are B



∃ 2x [x: A(x), B(x)]

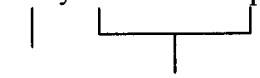
(12) All A are B



[x: A(x)] ∀ [: B(x)]

Geurts suggests that when children interpret a universally quantified sentence like (12), even though they construct a correct semantic form, they apply the weak parsing strategy. Therefore, when the sentence gets mapped onto its semantic form, the mapping is non-adult-like:

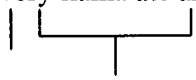
(13) Every llama ate an apple.



every [x, y: llama(x), apple(y), x ate y]

Since the child, according to this theory, knows lexical semantics of *every* and hence knows that it is a relational quantifier, but simply has difficulty with the mapping from form to meaning, her semantic representation actually looks like (14), where the domain of the quantifier is left underspecified:

(14) Every llama ate an apple.



[... : ...] <every> [x, y: llama(x), apple(y), x ate y]

Constructing such a representation, with an underdetermined domain of *every,* according to Geurts, causes the child to resort to pragmatic inferencing to determine the domain of quantification.

One pragmatic factor that determines the restriction of a quantifier domain in adult grammar is contextual salience - a contextually salient collection of individuals X is assigned to Q's domain. Another factor is information structure, namely the focus/background partition. Geurts points out that information structure interacts with syntax to narrow quantifier domains.

The material that in a neutral sentence is mapped onto the nuclear scope, in a sentence with focus is split: only the focused material becomes part of the nuclear scope, while the background material ends up being part of the restrictor regardless of its syntactic position. This is illustrated in the following examples, where sentences in (16) are paraphrases of (15), and the underlined part is the domain of *most*.[8]

(15a) Most Americans [visit a European country]$_F$ after they retire.
(16a) What most Americans who retire do is visit a European country.

(15b) Most Americans visit a European country [after they retire]$_F$
(16b) Most Americans who visit a European country, do so after they retire.

(15c) Most Americans visit a European country after they retire.
(16c) Most Americans are individuals who visit a European country after they retire.

According to Geurts, the interaction between salience and focus in narrowing quantifier domains is what is responsible for children's pattern of responses. He hypothesizes that children know this focus/salience interaction in determining quantifier domains. He claims that after children construct a representation (14) with its underdetermined domain, they then use salience and focus/background division within the nuclear scope determined by the context to construct the quantifier domain.

Thus, if a child assumes that a set of llamas is the most salient discourse entity, she would assume that it is the intended domain and that [x: llama(x)] is backgrounded and consequently the sentence is consistent with the adult interpretation. However, if the child focuses her attention on the set of apples, then the set of apples [y: apple(y)] is backgrounded and put in the restrictor giving semantic structure in (17):

(17) [y: apple(y)] <every> [x: llama(x), x ate y]

---

[8] These examples are an extension of the example provided by Geurts.

Hence, the peculiar 'symmetrical' errors – 'no' if every llama eats an apple but there are extra apples, 'yes' if every apple is being eaten by a llama but there is an extra llama, and a correct response if symmetry is observed.

The strength of this proposal is that it accounts not only for the classic spreading errors, but also for under-exhaustive errors, even if it does leave perfectionist errors unexplained.

The issue, however, is that despite its claims of children's full grammatical competence, this theory still attributes to children non-adult-like grammatical mechanisms. Essentially, it amounts to a claim very similar to the one made by earlier theories (that go back to Bucci (1978), Roeper and Matthei (1974) and Roeper and de Villiers (1993)), which explained the errors in terms of children's failure to find a correct restrictor for the universal quantifier. Unlike these earlier theories, for Geurts, 'spreading' of a quantifier domain is not due to children's incorrect lexical semantics of *every* or a failure to construct correct semantic representations, but to a combination of factors. Children are said to differ from adults in the following respects: (i) weak construal of a strong quantifier, resulting in non-canonical parsing, (ii) allowing the quantifier domain to be left underdetermined in the semantic form, and (iii), allowing pragmatic factors, namely the perceived salience of sets of individuals and information structure alone, instead of together with the syntactic structure, to determine the domain. While these pragmatic factors do play a role in figuring out quantifier domains for adults, in adult grammar they are used to narrow domain restrictions, not to determine which logical constituent the quantifier should range over, which is syntactically determined (notice that in all of the examples in (15)-(16), a difference in focus affected how the domain of *every* was narrowed, i.e. what relevant subset of the set of Americans was included in the domain). It did not change the restrictor set in the sense that there is no adult interpretation in which *every* would range over a set of European countries

instead of a set of Americans.

Furthermore, this theory attributes to children sophisticated pragmatic knowledge that needs to be corroborated by independent evidence. Thus, the claim that children rely on focus/background partitioning to restrict quantifier domain needs to be tested independently. Research conducted in that area is not without its own controversy. While some research has shown early knowledge of information structure (e.g. de Cat 2003), a substantial body of evidence suggests that children's knowledge of focus is not fully adult-like. Thus, studies by Szendröi (2003) for Dutch, Gennari et al. (2001), Gualmini, Maciukaite and Crain (2002), and Halbert, Crain, Shankweiler and Woodams (1995) for English showed that children are not sensitive to focus stress in comprehension task.[9] For example, in Szendröi (2003), children were tested with a truth-value judgment task on sentences like (18a), and (18b), which differed in the placement of focus stress:

(18a) X *only* verb Y *to* /Z/$_F$ (18b) X *only* verb /Y/$_F$ *to* Z,

In one condition, the stress was on the direct object, while in another on the indirect object. It was found that while children responded correctly (over 80 % correct) to sentences in which the stress was in the neutral position - on the indirect object, their performance was at chance level on the condition with a shifted focus stress. If we extend this preference for default stress to quantified sentences of the type we are investigating, it would suggest that we have no reason to assume that children would frequently choose to place focus on the subject (e.g. *every /llama/$_F$ ate an apple*). Furthermore, even if we assume that children can successfully place focus on the

[9] It is interesting that in production children's performance with marked stress is much better. In an experiment by Hornby and Hass (1970), discussed in Thornton and Wexler (1999), children between the ages of 3;8 and 4;6 were shown pictures and asked to describe them. The pictures were presented in pairs differing in only one element. Children stressed the contrastive element 80% of the times in their description of the second picture. As pointed out by Thornton and Wexler (1999) this asymmetry between comprehension and production in child language remains a mystery in subsequent studies.

subject and assign the object to be background, it is unclear how this would lead them to interpret the object as the restrictor for the subject quantifier. It would be a departure from adult grammar and hence needs to be corroborated by experimental data.

2.2. Full Semantic Competence Theories

2.2.1 Theory of Plausible Dissent

One of the most influential accounts attributing the errors under consideration to pragmatics is the theory of Plausible Dissent first proposed by Crain et. al (1996) and argued for in subsequent work by Gualmini et al. 2003, Meroni et al. 2000, etc. According to this view, the appearance of spreading errors in the contexts of visual asymmetry is purely coincidental. The errors are claimed to be not a response to the extra individual in the picture, as assumed by all other theories, but a reaction to the pragmatic infelicity of how the tests are typically administered. According to Crain et al., in order for a truth value judgment to be felicitous, the context has to meet certain pragmatic conditions, the relevant one for the issue at hand being the condition of Plausible Dissent. It requires for the addressee to be able to conceive of an alternative outcome in order to judge the proposition as true or false in a particular context. While adults and older children are capable of constructing alternative outcomes, according to Crain et al, young children lack the ability to do so and need the context to represent them explicitly. Since the picture-based tests do not suggest an alternative outcome (e.g. all boys are simply shown riding ponies, but it is not shown that some of them first had an option of riding an elephant or not riding at all), children cannot judge the truth or falsity of the sentence reliably. This theory would handle the 58%-rate of erroneous responses typically obtained with some groups of children (Philip's symmetrical group) as an indication of chance performance, while

the much higher 89%-rate obtained with other groups (Philip's perfectionist group) as a No-bias, when the no-answer is given by default if a correct answer cannot be determined.

The strength of this approach is that it successfully showed how experimental design affects children's performance, and that given optimal experimental conditions, children are capable of improving their performance to an error-free rate almost as low as that of adults; thus supporting the idea that the errors are not a matter of competence but of performance. Plausible Dissent Hypothesis also helps account for the fact that children are not limited to the non-adult responses. The condition of Plausible Dissent, however, has been criticized as unnecessary because researchers have been successful in improving children's performance by manipulations that did not involve the condition of Plausible Dissent. Thus, Drozd and van Loosbroek (2001) showed that the error rate can be reduced if the experimenter when presenting the context puts an emphasis on the restrictor set of the universal quantifier, regardless of how this emphasis is achieved. In their experiment, D&L had three experimental conditions: pointing, mentioning a property irrelevant to the test question or one relevant to it. An example of the irrelevant property condition can be the following: "Once there were these boys, one wore black shoes, another brown shoes, and the third one blue shoes. They saw these horses and wanted to ride them. One rode this horse, another this one, and the third this one." In the relevant property condition, the clarification might include mentioning the color of the horse each boy rode. Even though in every condition the context contained extra horses, the error rate was dramatically reduced across all three experimental conditions compared to the picture-only condition. D&L argued that since only the relevant-property condition satisfied the condition of Plausible Dissent, this condition cannot be a sufficient condition for children's successful interpretation of

universally quantified sentences. They argued that emphasizing the restrictor set was the relevant manipulation.

Phillip and Lynch (1999) found another way to reduce the rate of errors. They discovered that the error rate dropped from 25% to 10% when the picture contexts had the theme restrictor set changed from being in the foreground to being in the background. They argued that putting extra objects in the background made them less salient and hence irrelevant for the child, who then could ignore the asymmetry, and avoid giving incorrect false-judgments.

An experiment conducted by Crain et al (2000), disputed this assertion by showing that when the extra object was made highly salient concurrently with satisfying the condition of Plausible Dissent, the error rate was a low 11.4%.

Another interesting manipulation was done by Sugisaki and Isobe (2001). In this experiment, two groups of Japanese-speaking children were tested using picture contexts only. Group 1 was presented with pictures with one extra object (three cats each bouncing a ball plus one extra ball), Group 2 with pictures with at least 4 extra objects (three cat-ball pairs plus several extra balls). They found that by increasing the number of extra objects in the picture scenarios given to children, they were able to elicit a much higher percentage of adult-like responses: 37.5% correct for Group 1, in contrast to 87.5% for group 2.

Thus, if we look carefully at the complexity of the reported studies an interesting picture emerges. First observation is that rich linguistic context helps raise the rate of adult-like responses. Secondly, what seems to be playing a role in manipulating the rate of errors is a manipulation of the relative salience of the sets. This may explain why children's performance improves under such a wide variety of conditions – satisfying Plausible Dissent (Crain et. al 1994), creating picture-contexts with multiple rather than single extra individual (Sugisaki &

Isobe 2001), back-grounding of the extra individual (Philip and Lynch 2001), de-emphasizing the extra object (Gordon 1998), as well as emphasizing it (Crain et al 2000), and emphasizing the restrictor set of the universal quantifier (Drozd and van Loosbroek 2001).

I will argue that the Full Competence Approach is correct in insisting that children's semantics for universal quantification is fully adult-like; however, I will reject the assumption that children's apparent reaction to the asymmetry in the visual context is merely an experimental artifact. Nor will I rely on Plausible Dissent as the primary explanation of differential results obtained by various researchers. I will suggest instead that satisfying the Condition of Plausible Dissent always enriches the context, and the richness of context achieved either linguistically or visually is the best predictor of children's success with quantified sentences across languages and experimental conditions. I will discuss the role of context in my theory of children's knowledge of quantification in more detail in chapter 3.

2.2.2 Presuppositionality Accounts

If we look at more recent theories of children's grammar of universal quantification, such as Drozd and van Loosbroek (2004) and Philip (2004), we would see a shift away from positing any non-adult-like grammatical properties. Thus, D&L (2004) developed the Presuppositionality Account proposal. According to this hypothesis, even though the child responding with symmetrical errors to a universally quantified sentence is well aware that *every* is presuppositional, during the process of verifying this sentence, he is overwhelmed by the task of relating the domain presupposition of *every* to the truth conditions of the nuclear scope and ends up with "no presupposed set in mind to compare to the nuclear scope" (Drozd and van Loosbroek 2004: 5). As a result, she verifies '*Every* X *is* verb_{ing} *a* Y' in such a way as if it had

the meaning similar to '*Many* X *are* $verb_{ing}$a Y.'[10] This allows the Presuppositionality Account to provide an explanation of both under-exhaustive errors and both types of symmetry errors. It may also, however, predict errors of judging a sentence false if the complete exhaustive relation holds between subject and object sets.[11]

Philip 2004 developed another version of this account as The Relevance Account (RA). The RA attributes symmetry responses to partial knowledge of a pragmatic rule hypothesized to constrain adult verification of universally quantified sentences. This theory, like Geurts 2003 and Drozd and van Loosbroek 2004, draws a connection between symmetrical errors and the well-known property of quantifiers to have their domains restricted contextually. To put it in a simplified way, a child who judges the sentence like 'Every dog is licking a cat' false in the context of an extra cat does so because he imagines an unseen fourth dog—i.e. adds a cognitive object representing this unseen object to her context model—and defines the domain restriction for the universal quantifier as containing not just the dogs seen in the picture, but also the absent dog (failing to lick a cat). This is why he explains his judgment of falsity by pointing out that the cat in the tree is not being licked (by a dog). To explain why the child takes an unseen object to be salient, Philip uses the following illustration (fig 1):

A A A A A A

B B B B B

FIGURE 1 Illustration of a symmetrical pattern recognition.

The claim is that when looking at an asymmetrical pattern, as in figure 1, one cannot help but

---

[10] I understand this claim to mean that when children verify the truth-value of universally quantified sentences without any presupposed domain restriction, they may include unseen individuals into the domain or possibly exclude some of the visible individuals from the domain. In all other respects, the semantics of *every* for them is fully adult-like. The meaning of the universal appears to be similar to 'many' from the perspective of the experimenter, who does not have access to the domain restriction constructed by the child. For the child himself, the meaning remains that of a universal quantifier.

[11] Such errors were indeed obtained by Altreuter and de Villiers (2005). It is, however, generally assumed that children's knowledge of universal quantification (or their partial knowledge) involves correct 'true' judgments in this situation.

notice that a letter B seems to be missing. This intuition is the output of an unconscious process of symmetrical pattern recognition. According to the RA, it is this universal cognitive ability to recognize symmetry (and the absence of it) that causes the child to imagine an unseen object companion for the visible extra object in the paradigmatic case of exhaustive pairing. Symmetrical pattern recognition makes an unseen object salient for the child.

The RA accounts only for what Philip calls exhaustive pairing (what I have been referring to as symmetry errors) and contends that under-exhaustive and perfectionist (or 'bunny-spreading') errors have to and should be handled with a different mechanism. Until this claim of independence of the classic spreading errors from the perfectionist and under-exhaustive errors gets independent empirical support, a theory able to account for all three types of errors should be preferred as more parsimonious.

The RA account has another empirical weakness. As Philip acknowledges, it does not straightforwardly predict his observation that Dutch children showing exhaustive pairing for universally quantified subjects, often find sentences analogous to (19a) false in contexts like (19b) and true in contexts like (19c), the exact the opposite of adult judgments (Philip 2004).

(19a) An elephant is carrying every dog.
(19b) 1 elephant carrying 3 dogs + 2 extra elephants. adult-true / child-false
(19c) 3 elephants each carrying a different dog. adult-false / child-true

It shows that Dutch children, in contrast to adults, do not require a specific reading of an indefinite subject in sentences analogous to (19a). Similar data have been obtained from child Chinese (Lee 1986, Chien 1994; cf. also Sano 2003 for Japanese). It is, however, not the case that children, English as well as Dutch, generally have difficulty assigning specific readings to indefinite subjects in sentences like (19a). In fact, Musolino (1998) reports that English children have no difficulty doing this (cf. Philip and Botschuijver 2004). Likewise, Gavarro and Escobar

(2002) reported a high adult-like rate of responses (94% correct for the youngest group) on sentences analogous to (19a) for both accepting the surface and rejecting the inverse scope readings. Thus, we need to explain the discrepancy between the Dutch/Chinese and Spanish/Catalan results, as well as reconcile the claim of children's full semantic competence with the failure of the Dutch children to choose the correct interpretation for (19a).

The Presuppositionality accounts acknowledge that contextual asymmetry may be leading to children's errors not because their grammar of universal quantification calls for symmetry, but because of certain difficulty children have with establishing correct contextual domain restrictions for the universal quantifier – either because calculating the domain presupposition exerts too heavy a processing load on the child (D&L 2004) or because the visual asymmetry activates children's assumed symmetrical pattern recognition mechanism causing them to add unseen individuals into the domain (Philip 2004).

I will present evidence that this line of reasoning is on the right track and argue that non-adult-like calculations of domain presuppositions is the reason for the observed errors. I will also argue that this problem is not limited to the universal quantifier, but may be a more general problem. It is especially acute with indefinites, which I will discuss in chapter 3. I will show that positing incorrect domain restrictions for the indefinites may account for "classic spreading", while the analogous miscalculation of the domain restriction for *every* leads to under-exhaustive errors.

According to my proposal, the contextual asymmetry is responsible for the errors by providing pragmatic clues for how to restrict the domains of quantifiers, particularly when the experimental design offers no linguistic context to supply that information. An asymmetric picture misleads the child by suggesting that not all of the individuals are relevant for the truth-

value of the sentence; i.e. it suggests that the picture consists of two sets – a set of pairs and a set of single individuals, which need to be looked at separately. Natural discourse or an experimental design providing rich linguistic context would supply means to avoid such errors. Experimental conditions that use minimally informative pictures leave it up to the child to figure out how narrowly the domains should be construed and which of the two sets – boys riding ponies or the single pony - is what the sentence is about. A single forgrounded pony is likely to be more perceptually salient than a number of identical boy/pony pairs. Thus, asymmetry creates unequal visual salience of individuals in the picture, and is read by the child as a clue to what experimenter is asking about. In other words, the problem is not that asymmetric pictures violate the truth conditions for a universally quantified sentences, but that the visual clues are incorrectly used by the child compensates for the contextual paucity.

2.3 Why look for a new explanation?

As we have seen from the previous discussion, researchers had uncovered certain systematic ways in which children differ from adults in their interpretation of universally quantified sentences revealed under certain experimental conditions. The errors in question are very robust and have been replicated in a variety of languages (e.g. for Dutch in Philip and Verrips 1994, Philip & Coopmans 1995, Drozd and van Loosbroek 1999, Turkish in Freeman and Stedmon 1986, Japanese in Takahashi 1991, Sugisaki and Isobe 2001, Catalan Philip 1995, Gavarro and Escobar 2002, Chinese in Chien and Wexler 1989, and French in Inhelder and Piaget 1959, etc.). Such well-established behavior presents a challenge to the approach treating the 'symmetry errors' as nothing more than an experimental artifact, a mere illusion, which arises as a function of pragmatic infelicity and is purely coincidental with the conditions of

asymmetry in the sentence presentation. Given how systematic and predictable these findings are, they strongly suggest that the children do indeed respond in a non-adult-like-like way to asymmetries in the context in which the target sentences are presented. There are, however, strong reasons to doubt the explanations placing this 'symmetry reflex' within child grammar of universal quantification.

Let's summarize the objections to this approach.[12] First, any account of child responses has to explain why at any point in development children are not limited to the non-adult-like responses. Thus, Philip 1995 reports that the group he calls 'the symmetry children' gave 'symmetry' responses only about 57% of the time. The "grammatical symmetry" theories would be forced to conclude that children's grammar contains two mechanisms of universal quantification: adult-like and non-adult-like. This implies a developmental progression from a grammar allowing two options (the universal quantifier optionally quantifying over events or over individuals) to a grammar in which only one of these options is permitted. Such developmental path, in my view, is not easily accommodated within straightforward Piaget-style maturational stages, which usually progress from simpler, more primitive representations to more complex ones, and not from two equally complex options to one. In my view, not being able to appeal to a greater simplicity of the earlier stage as compared to the later stage, or to point out precisely what deficit causes children to have the non-adult grammatical option, as was done by Inhelder and Piaget, weakens the stage approach by making it ad hoc and less explanatory.

Another difficulty for the "grammatical symmetry" approaches is to justify the existence of pre-adult-like structures they posit, i.e. structures randomly different from those found in adult

[12] See Crain 2002 for additional arguments against this view from the perspective of the theory of learnability.

language.[13] Such structures are very unusual in children's spontaneous production and quite rare in experimental data. In those rare instances when children do exhibit unusual (for their target language) structures, the error is not random, but constitutes a grammatical option for some adult-language. Thus, the error of inserting a medial wh-phrase in long-distance questions reported for English–speaking children, is a grammatical option found in certain German dialects and is not confined entirely to child language (Thornton 1990). Even though it may be necessary to posit a child's grammar with arbitrary formal properties, we have to be cautious and resort to that explanation only if other alternatives fail, as discussed in chapter 1. This principle, known as the Continuity assumption, is motivated by the considerations of learnability, i.e. the need to explain how a child would progress from a non-adult-like grammar to a full adult-like competence given limitations of the input (Pinker 1984).[14] Of course, it is a logical possibility that for acquiring adult-like competence no particular properties are required from the input or child's internal grammar acquisition mechanism because adult-like properties simply mature on a genetically pre-determined schedule. However, since this theory is unconstrained and ad hoc (it would allow us to posit a grammar for any linguistic behavior we may encounter in children and claim that it goes away by itself through maturation), I agree with those who insist that we resort to a maturational account only after we rule out alternative explanations.[15]

---

[13] Even though event quantification is part of adult grammar, adopting it as an explanation for children's errors would commit us to accepting non-adult-like mechanisms into child grammar, namely odd restrictors formed by conjunction of syntactic elements that do not form a syntactic constituent or even simply allowing children to interpret a determiner as an adverb disregarding syntactic structure. Evidence that children know syntax of DPs comes from studies of spontaneous production that show that English-speaking children as early as 18 months of age almost never make grammatical errors in the use of DPs (Abu-Akel and Bailey 2000).

[14] As Pinker put it, "...for us to take a characterization of a developmental state seriously, there must be a way into the developmental state and also a way out of it..." This consideration imposes a severe constraint on descriptions of the rule systems underlying child language." (p. 6)

[15] Another observation about the symmetrical errors that, in my view, is not easily compatible with a grammatical stage approach is their high variability across children, wide age range, and gradual protracted decline. Q-spreading has been reported for children between ages 3 and 7-8, while groups of children as young as 3 have been reported to be fully adult-like (Philip 1995). A genetically predetermined maturational schedule in the

The most compelling reasons for questioning the grammatical symmetry are of empirical nature. One observation to be explained is that the symmetrical errors appear to be robust only in an experimental setting and only in comprehension. Naturalistic studies of children's use of determiners report very early competence (e.g. Abu-Akel and Bailey 2000). As far as I know, no elicited production studies have produced corroborating evidence for q-spreading.[16] Furthermore, the available experimental data clearly show that the symmetrical errors are limited for the most part to certain types of experimental conditions. The basic finding of the 'q-spreading' literature is that a high percentage of non-adult-like responses is found exclusively in tests in which context is given visually (as picture scenarios), where no or little linguistic context is given. If the symmetry requirement were a result of a grammatical mechanism, we would expect it to be evident under other experimental conditions as well. There should be no reason to expect the symmetry requirement to disappear simply because the context is enriched if it were indeed part of grammar. Furthermore, certain quite predictable ways of manipulating the rate of error experimentally has been discovered, as discussed in section 2.2.1. Thus, not only enriching the context has helped reduce children's rate of errors, but also increasing the number of asymmetry-creating individuals, deemphasizing the extra individuals in some way, or conversely emphasizing them and hence making it clear that these individuals are 'extra', and hence irrelevant. I would argue that since the way of introducing the universe of discourse for the truth value judgment matters so much in how well children perform on this task, we should look at

---

population of typically developing children may be expected to be more uniform and perhaps represent a more-or-less abrupt shift in performance at the point when the child progresses into the next developmental stage.

[16] Even though Altreuter and de Villiers 2005 claim to have produced such evidence, their data contains responses of children failing to produce universally quantified sentences in response to pictures with asymmetrical sets, who substitute them with other types of sentences. This is less compelling than actual non-adult-like 'spreading' production data would be. In order to elicit such data, it may be necessary to use negated universally quantified sentences like 'Every boy is not riding a horse' in a typical 'spreading' context – e.g. with an extra horse. This should be only minimally different from a typical truth value judgment task in which a child gives a 'no-response' to the affirmative sentence.

their errors as a function of the pragmatic conditions of the test, and hence seek an explanation in their pragmatic skills, instead of treating them as a function of incorrect grammar.

Another empirical challenge to the "grammatical symmetry" theories is that they have certain implications that were challenged by subsequent research. Thus, they postulate a mismatch between the syntactic and semantic structure on the non-adult readings, when the quantifier is not restricted by the NP that forms a syntactic constituent with it, but by a disjunction between the set denoted by the subject and object NPs. This entails a loss of distinction between the two syntactic arguments of the universal quantifier. Consequently, sentences that differ in the position of the universal quantifier (it being in the subject vs. the object position) would have essentially the same logical form and, as long as the requirement for symmetry is not met, both sentences should be judged by the child as false at similar rates. This claim has been disputed.

Meroni et al (2002) showed that children possess a knowledge of the asymmetry between internal and external arguments of the universal quantifier with respect to the property of Downward Entailment (DE). In adult language, the two arguments of the universal quantifier *every* differ with respect to this property. The quantifier is downward entailing only in its internal but not in its external argument, i.e. an inference from a set to its subset is licensed for the internal argument of the universal quantifier (from *boys who rode an elephant* to *boys who rode a wild elephant*), as shown in (20), but it fails for not the external argument (from *carried a flag* to *carried an orange flag*) as shown in (21)[17]:

(20) Every [boy who rode an elephant] [carried a flag]=1 →
Every [boy who rode a wild elephant] [carried a flag] =1.

(21) Every [boy who rode an elephant] [carried a flag] =1 ≠>
Every [boy who rode an elephant], [carried an orange flag]

[17] The following examples are somewhat modified from the original examples used by Meroni et. al..

One consequence of the asymmetry in DE between the two logical arguments of *every* is that *or* can receive either an exclusive disjunction interpretation or a conjunctive interpretation in the internal argument of *every*, but only the exclusive disjunctive interpretation in its external argument.

(22) Every [boy who rode an elephant or a zebra] [carried a flag] ⇔ Every boy who rode an elephant carried a flag and every boy who rode a zebra carried a flag.

(23) Every [boy who rode an elephant] [carried a flag or a balloon] <≠> Every boy who rode an elephant carried a flag and every boy who rode an elephant carried a balloon'.

Meroni et al. 2002 tested whether English-speaking children know the difference between the internal and external arguments of *every* with respect to the contrast in licensing of conjunctive *or*. They report that children know that *or* can be used conjunctively in the internal argument of *every*, but only as exclusive *or* in the external argument. [18],[19]

Finally, certain skepticism with respect to the grammatical account of the 'symmetry reflex' is due to the findings that children make errors unpredicted by 'q-spreading' – types of accounts. After all, a strength of a grammatical account is judged by its ability to make testable predictions about what types of errors should occur and what should be ruled out. In the case of children's quantificational errors, no grammatical theory accounts for the full range of errors and does not predict a full range of errors present in the experimental data. Thus, children do not only make errors when they seem to 'spread' the domain of the universal to unexpected parts of the sentence (e.g. when they incorrectly judge a sentence like (0) false in the context of a picture containing an extra pony), and which may suggest that 'every' ranges not only over boys, but

---

[18] The knowledge of this property was also tested by Gualmini et al, 2001, who tested whether children allow conjunctive use of *or* in the internal argument of *every*. They report children's competence in this respect.

[19] Further evidence for children's sensitivity to the syntactic position of the quantifier, comes from a report that the rate of children's errors is affected by whether the universal is in the subject (right-spreading error) or object (left-spreading error) position. Thus, H-K Kang (2002) reports a significant difference between the rate of the right-spreading (51%) and the left-spreading error (73%) in English-speaking children.

also over ponies. They may also judge the sentence false if the picture contains any other object unmentioned in the test sentence (e.g. a girl riding a camel). Furthermore, they may also incorrectly judge the sentence true if the picture contains an extra non-pony-riding boy (a so-called under-exhaustive error). Some experimenters even report a substantial number of children who judge a universally quantified sentence false despite a perfect symmetry between boys and ponies and no extra unmentioned individuals (Altreuter and de Villiers 2005). Such wide range of errors suggests that the errors stem from something other than the formal semantics of *every*, which would have to be quite unconventional in order to cover the range of readings corresponding to the types of quantification errors reported in the literature.

After carefully considering the experimental results reported in previous research, and noticing that the rate of errors can be consistently manipulated through the experimental technique, I suggest that that the errors are not due to a lack of certain semantic knowledge (or cognitive competence underlying it), but are task-specific effects. It is desirable to find a unified explanation to the various attested results able to handle the various experimental effects discussed above. The goal is to reconcile the claim that children possess the knowledge of the semantics of the universal and are able to restrict the quantifier correctly with the observations that their performance is in some sense vulnerable and can be easily sabotaged. The question we have to answer is where the source of this vulnerability lies. We will try to answer this question in the most parsimonious way without adding any linguistic mechanisms to the standard grammar of quantification of English. The theory proposed in the next section will not dismiss the 'symmetry-reflex' data and will treat it as a real phenomenon in children's performance. Even though I will continue to use the term 'symmetry reflex' and consider it a real phenomenon and a descriptively useful term, I will not consider it a grammatical phenomenon. I will maintain

that children's 'every' means exactly what it means for adults. Lack of symmetry does indeed triggers errors, but not because child's grammar requires contextual symmetry. I will present new experimental data showing that the 'symmetry reflex' goes beyond the universal quantification errors, but also shows up in sentences with the negative determiner (e.g. no boy kissed a girl). I will also present new data showing that a grammatical account is unnecessary for accounting for the symmetrical errors. In this proposal, the source of errors is pragmatic and is closely related to the children's weakness in 'reading other minds'.

Chapter 3. Theoretical Preliminaries

1. What is the role of the lonely donkey?

1.1 Hypothesis Recapped

In chapter 1, I have proposed to look at the ambiguity of sentences like (1), which is due to the interaction of the two quantifiers, the universal and the indefinite, as the cause for children's quantification errors.

(1) Every boy rode a pony.

I suggested that the crucial source of difficulty lies not in the interpretation of the universal, but in the "free" nature of indefinites with regards to how their domain restriction can be set - in addition to being permitted to have a domain delineated as a singleton or a non-singleton set without any morphological clues to which option has been chosen, indefinites permit an information gap between the speaker and hearer in having access to the domain (Schwarzschild 2002, Kratzer 1998). I would like to argue that children have the semantic knowledge of both the universal and the indefinite; hence, from early on they have the knowledge of the ambiguity. However, they are less good than adults at resolving it in a minimal context. What appear to be 'spreading' errors are not grammatical errors. They are a manifestation of a strategy children use to resolve the sentence ambiguity by restricting the domain of the indefinite to a singleton set if the situation contains a salient single object. The presence of the latter, which may be taken as a 'referent' for the 'wide scope' indefinite, often leads the child to assign 'wide' scope to the indefinite when adults wouldn't. In order to understand how this works, let's take a closer look at the semantics of indefinites.

1.2 Quantifier Domain Restrictions

As has been pointed out, an important semantic feature of quantifiers is that they range over **relevant** individuals. Thus, (1) does not say that every boy in the world rode a pony. If it did, there would be very few situations in which (1) were true.

It has been argued that this notion of relevance in quantification is part of the semantics of quantification and hence has to be part of the formal semantics of quantifiers. Quantifier domain restrictions have been posited as the linguistic mechanism of expressing "relevant" (Westerstål 1984, von Fintel 1995, Marti 2003). Under one common view of how they work, they are expressed as silent pronoun-like elements (e.g. a sister to the common noun phrase argument of the quantifier) in the semantic representation, a common notation for which is shown in (2) and its abbreviated semantics given in (3):

(2) Every boy_C rode a pony.
↑
*context supplies its value: e.g. "**those who came to Sam's birthday party**"*

(3) Every [C & boy][rode a pony] or
Every [λx.C(x) & boy(x)] [λx.rode-a-pony(x)]

This context-dependency is a property that indefinites share with other quantifiers. If this is correct, interpreting (2) involves not only choosing relevant boys (the value of C), but also relevant ponies (the value of C'):

(4) ∀x [x is in C & x is a boy →∃y [y is in C' & y is a pony & x rode y]] ,
↑
*context: **"the ponies Sam's parents hired for the birthday party"***
or narrowed to a singleton set: ***"the white pony"***

Furthermore, choosing one relevant pony instead of multiple ponies has important consequences for the relative semantic scope of the indefinite and the universal quantifier: it creates an interpretation identical to the one in which the sentence has the

inverse scope interpretation[1]: If C'= $\{P_1, P_2, P_3...\}$, this yields a non-specific indefinite and allows co-variation of boys and ponies. If C'= $\{P_1\}$, it yields a specific indefinite, truth-conditionally identical to the wide scope indefinite resulting in (4) being equivalent to (5):

(5) $\exists$y [y is in C' & y is a pony and $\forall$x [x is in C & x is a boy and x rode y]]

What this demonstrates is that if the restriction of an indefinite is true of exactly one individual (what Schwarzschild 2001 calls "singleton indefinites"), it behaves logically like a referential noun phrase. In such a case, in order for the sentence to be true, all of the boys are required to ride the pony included as the singleton restriction.

Indefinites are notorious for their 'irregular' behavior: their seemingly varying quantificational force, when they appear to be existential quantifiers in certain syntactic contexts but not in others (Lewis 1975, Kamp 1981, Heim 1982, among others), and their ability to take island-free wide scope (Fodor and Sag 1982, Reinhart 1997, Kratzer 1998, among others). There have been various approaches proposed to account for the behavior of indefinites, reviewing all of which is outside the goals of this dissertation. Although multiple theories of indefinites may be able to handle the child data I will consider here, for concreteness I will adopt one of them, namely Schwarzschild 2001, who argues that indefinites are unambiguous existential quantifiers with their scope derived in a regular way. According to this proposal, the exceptional wide scope of indefinites can be derived simply by manipulating the cardinality of their domain restrictions. In the remainder of this chapter, I will review this theory and discuss how this approach sheds light on q-spreading as well as children non-adult-like behavior with respect to quantifier scope.

[1] I adopt the presuppositional theory of 'every' (i.e. it presupposes a non-empty set of boys in (3)).

1.3 Deriving the Scope of Indefinites

It has long been assumed that ambiguities observed with sentences with multiple scope-bearing elements, such as quantifiers, are derived through a covert movement operation (QR), which creates operator-variable structures by moving and adjoining the operators to the left of clausal boundaries (e.g. adjoining to IP), with the operator binding the variables from its raised position (Chomsky 1976, May 1977, 1985):

(6) [IP ... QP ...] → [QP [IP ... t ...]]

Thus, a sentence with an indefinite and universally –quantified NPs has two readings – with either the universal or the indefinite taking wide scope, the readings paraphrased in (8) and (9) and whose abbreviated LFs are given in (10) and (11).

(7) A boy tasted every dish.

(8) For every x, such that x is a boy, x tasted a dish.
(9) There is an x such that x is a boy and x tasted every dish.

(10) [Every dish$_2$ [a boy$_1$ [IP t_1 ... t_2...]]]
(11) [a boy$_1$ [Every dish$_2$ [IP t_1 ... t_2...]]]

An important argument for the movement theory of scope comes from the data showing that QR obeys locality constraints on movement, such as the "island constraints", barring extraction from strong islands.[2] Thus, Hornstein 1984 observed that while the universally quantified NP can take wide scope from the embedded subject position in the

[2] Among the arguments for QR are not only that movement creates a transparent structure off which to read the relative quantifier scope, but also the considerations of interpretability (to avoid type mismatch between the transitive verb and the quantifier as its internal argument), deriving inversely-linked readings: structures in which the more embedded QP takes wide scope (e.g. *One apple in every basket is rotten*), ACD constructions, quantifiers that bind pronouns, etc. (see Heim and Kratzer 1998 for a review).

ECM construction, as in (12), it takes obligatory narrow scope in a tensed clausal complement, as in (13).

(12) A professor expects every student to pass the exam. (a reading with co-varying between professors and students).
(13) A professor expects that every student will pass the exam. (no co-variation)

It has also been observed that the movement theory of scope runs into problem with indefinites, which do not obey island constraints. Consider the example from Fodor and Sag (1982) given in (14), which has a reading in which the indefinite takes wide scope out of an adjunct island:

(14) If a relative of mine from Texas dies in a fire, I will inherit a house.

One can imagine that the speaker has multiple relatives from Texas. If this is the case, two interpretations are possible. Under one reading, a death in a fire of any one of the speaker's relatives from Texas would result in the speaker's inheriting a house. Under another scenario, one specific relative must die in a fire in order for the speaker to inherit a house. Accounting for this reading with a wide scope indefinite construed as the existential quantifier is problematic since such scope taking would be exceptional (moving out of a syntactic island) and is disallowed for other quantifiers. To see the contrast, we can look at a minimally different sentence, e.g. with the universal quantifier instead of an indefinite in an adjunct clause, as in (15), which is unambiguous: 'every' is not permitted to scope above the conditional (the reading under which every relative of the speaker's is such that the speaker will inherit a house in case that relative dies in a fire). The only available reading is the one in which death of all of the relatives in a fire is required for him to inherit a house.

(15) If every relative of mine dies in a fire, I will inherit a house.

An even more serious problem is presented by sentences with indefinites that have intermediate scope readings, as in (16):

(16) Every author here despises every publisher who would not publish a book that was deemed pornographic.

(17) Possible readings:

a. $\forall$x (author x) >$\forall$y (publisher y) > $\exists$z (book z): For every x, if x is an author, then x despises every publisher y who refused to publish some book z or another that was deemed pornographic.
b. $\exists$z (book z) >$\forall$x (author x) >$\forall$y (publisher y): There is a z, such that z is a book that was deemed pornographic and for every x, x is an author, x despises every y, y is a publisher who refused to publish z.
c. $\forall$x (author x) > $\exists$z (book z) >$\forall$y (publisher y): For each x, x is an author, there is a book z that was deemed pornographic – possibly different books for different authors – such that x despises every publisher y who would not publish z.

The existence of the intermediate indefinite readings as in (17c) for sentences like (16) presents a problem for both the 'naïve' quantificational view of indefinites (Russell 1919) because, as discussed above, unlike other quantifiers in this configuration, they do not respect islands in their scope-taking, as well as the ambiguity/bifurcation view (Fodor and Sag 1982, Hornstein 1995). The latter theories can handle the widest-scope reading in (17b). However, the intermediate reading (first discussed by King 1988) is problematic for this approach.

The analysis of indefinites proposed by Schwarzschild (2002) successfully solves the problem of exceptional scope including intermediate scope using a mechanism that affects domain restrictions of all quantifiers.[3] It also provides a mechanism that can be used to explain q-spreading behavior of children, as I will show in section 1.5.

[3] Another solution to this problem suggests that exceptional wide-scope indefinites are interpreted as choice functions (Kratzer 1998, Winter 2001, Reinhart 1997, etc.). A choice function takes as its argument the set of individuals satisfying the descriptive content of the indefinite NP and returns an element from this set. This approach is also compatible with the analysis of q-spreading I am proposing. The problems with the choice function approach were discussed in Chierchia 2001, Schwartz 2001, Geurts 2000, but these objections do not bear on the data discussed here and discussing them is beyond the scope of this dissertation.

1.4 Singleton Indefinites

Schwarzschild's insight was that the specific/non-specific ambiguity of the indefinites is not lexical but can be derived simply by adjusting the size of their domain restriction, which, according to Schwarzschild, can consist of a set of any cardinality (except zero) including being a singleton set. This insight allows us to maintain a uniform analysis, according to which indefinites are unambiguously existential quantifiers and their syntactic scope is determined and constrained in a regular way. However, if the domain restriction of the indefinite is contextually narrowed to a singleton set, their syntactic scope may get masked. Under these circumstances, the indefinite appears to have wide scope without applying covert movement and regardless of its surface c-command domain, a situation Schwarzschild considers to be of scope neutralization.[4]

According to Schwarzschild's proposal, the reason why exceptional 'wide scope' is permitted with indefinites is because unlike other quantifiers, they freely allow, although do not limit their domain restrictions to singleton sets. In order to have this latter reading, the context must provide a unique relevant individual that can be included in the singleton restriction for the domain of the indefinite. If this is the case, then the surface scope reading would have truth conditions identical to those in which the indefinite has syntactic wide scope, a configuration which then becomes unnecessary for explaining the 'specific' reading. Hence, positing any exceptional scope taking capabilities of the indefinites becomes unnecessary, including the lexical quantificational/referential ambiguity.

[4] Scope of Q_1 is neutralized relative to Q_2 if Q_1 takes scope over Q_2, but if their relative scope were reversed, truth conditions would not be affected.

This elegant solution to the puzzle of indefinites allows us to account for all of the exceptional scope readings, including the intermediate scope, without any additional syntactic or semantic mechanisms. If we correctly construct the restricted domain for the indefinite as a singleton set and fill in the implicit domain-narrowing restriction, we get the desired truth-conditions without having to move the indefinite out of an island, hence avoiding the need to account for its apparent exceptional scope. According to this proposal, different readings that we saw in (17) are due to a differently constructed domain of the indefinite. All of the readings therefore are narrow-scope readings, but for two of them their scope is masked by the fact that the indefinite is restricted with a singleton set.

In order to illustrate how this works more explicitly, let's suppose that the set of relevant authors includes three individuals as shown in (18). Let's also suppose that for each of these individuals x there is a distinct single book (e.g. a book that x wrote) that was deemed pornographic, and that each author x despises every publisher who refused to publish that book (i.e. the domain restriction of the indefinite contains a bound variable). The formula in (19) shows that even though the indefinite occupies the lowest position, we get the desired co-variation between authors and books seen in the 'intermediate' reading.

(18) Authors: {Lawrence, Miller, Nabokov}, Books: C_L={Lady Chatterley's Lover} C_M={Tropic of Cancer} C_N={Lolita}, Publishers: …

(19) $\forall x$ [author(x) $\rightarrow \forall y$ [publisher(y) such that $\exists z$ [z $\in C_x$ and z is a book and z was deemed pornographic and y refused to publish z] $\rightarrow$ x despises y]] where C_x={z: z is a unique contextually relevant book **x had written** and z was deemed pornographic}

Another reading, the regular narrow scope reading, is derived with the domain of the indefinite consisting of multiple books, as shown in (20):

(20) Authors: {Lawrence, Miller, Nabokov}, Books: C={Lady Chatterley's Lover, Lolita, Tropic of Cancer, ...}

(21) $\forall x$ [author(x) $\rightarrow \forall y$ [publisher(y) such that $\exists z$ [z $\in$C and z is a book and z was deemed pornographic and y refused to publish z] $\rightarrow$ x despises y]] where C={z: z is a book that has been deemed pornographic};

Here the domain may contain all books that have ever been deemed pornographic, or it may be contextually narrowed to a subset containing multiple relevant ones.

Finally, the widest scope reading can be derived in the same way as a narrow scope indefinite, but with its domain restricted to a single book:

(22) Authors: {Lawrence, Miller, Nabokov}, Books: C ={Lolita}

(23) $\forall x$ [author(x) $\rightarrow \forall y$ [publisher(y) such that $\exists z$ [z $\in$C and z is a book and z was deemed pornographic and y refused to publish z] $\rightarrow$ x despises y]] where C={Lolita}

In this example, the domain also contains an implicit restriction. The implicit restrictions in this case, as with any other singleton indefinite, may be asymmetrically available to the speaker and not to the listener and may contain any conceivable property narrowing the set to a unique book. What this means is that when uttering a sentence with an indefinite, the speaker has no expectation that the listener should be able to reconstruct from the context the indefinite's domain restriction with any specificity (beyond figuring out that it contains a single individual, whose identity is known to the speaker).[5] In other words, the speaker may felicitously utter the sentence with a single individual in mind and have no expectation that the listener would infer from context what individual it is.

[5] This information gap between the speaker and the listener may not be simply permitted, but required. If the speaker and the listener both share the knowledge of which individual is in the domain restriction, then a definite NP is required. For example, if the identity of a rich uncle from Texas is known to the listener, then (14) becomes odd, and a definite "the relative of mine from Texas' is required.

Thus, Schwarzschild observed that indefinite 'wide' and 'intermediate' scope readings for (16) can be derived with the same syntactic scope configuration as the narrow scope reading. The difference between them is due solely to a differently constructed domain for the indefinite, the value of which is determined contextually. This property of indefinites, when their ambiguity is accounted for not with a syntactic, but with a contextual manipulation, is what makes doubly quantified sentences with indefinites difficult for children. As I will argue, children have complete adult-like knowledge of all grammatical rules (syntactic and semantic) that regulate quantifier meaning including their scope taking. On the other hand, as I suggested in chapter 1, children have difficulty managing contextual information and, in particular, correctly gauging the epistemic effects of context on other interlocutors (i.e. judging correctly what a particular context would make known/obvious or salient to others present in the situation – a skill underpinned by a non-linguistic, ToM-realted, ability to reason about other minds). Therefore, it is not surprising that children would frequently fail to construct the target interpretation for sentences with indefinites, particularly in those situations when the context given to them is less than optimal.

According to Schwarzschild's approach, any indefinite can be "singleton"; hence, in an out-of-the-blue context they are ambiguous with respect to their specificity. Furthermore, there exists an information gap between the speaker and the listener with respect to the membership of the set constituting its domain (i.e. indefinites are vague from the listener's perspective). This vagueness, in addition to their ambiguity between singleton and non-singleton readings, is what I propose makes doubly quantified sentences with indefinites hard for children. The interpretive errors of quantification are

caused by children's weakness in precisely this area – managing information gaps between them and others (e.g. assessing others' knowledge and providing the right amount of information).

To sum up, in my solution of the puzzle presented by children's errors on universally quantified sentences, I propose that by focusing all our attention on the universal, we had overlooked the role of indefinites with their ability to change the meaning of a universally quantified sentence with a contextual manipulation supplying a unique individual for its domain restriction. In order to interpret a sentence with a universally quantified subject and indefinite object, one not only has to identify the relevant individuals to be included in the subject set, but also to make a decision as to what subset of individuals in the object set is relevant and narrow the domain restrictions accordingly. This task would be especially difficult if the target sentence is presented without sufficient linguistic context.

1.5 Implications for 'quantifier-spreading'

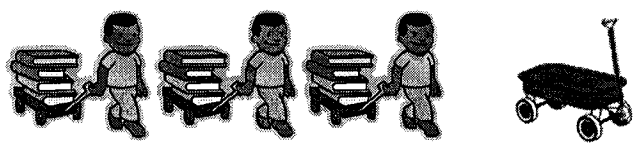
The observation about the role of the domain restriction in the interpretation of sentences with indefinites has important implications for the question we are considering. I have reviewed arguments for the importance of the domain restrictions for quantifiers in general and indefinites in particular. I have established that in order to evaluate a sentence containing an indefinite NP with respect to its truth conditions, we need to know not only the relative syntactic scope of quantifiers, but also the value of the restricted domain for both the universal and the indefinite.

The latter presents difficulty because any indefinite can in principle have a domain restriction of the cardinality of one. The difficulty is magnified by allowing the

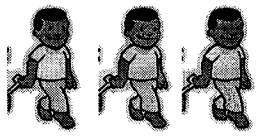
restriction to be expressed implicitly, making it asymmetrically available to the speaker, but not the listener. This creates a possibility of interpreting any indefinite as singleton or non-singleton and is a reason behind the intuition for the specific/non-specific ambiguity of indefinites in an out-of-the-blue context. Importantly, this suggests that a sentence that contains an indefinite NP presented in a relatively poor context is ambiguous regardless of the availability of a genuine inverse scope, and a relatively rich context is necessary for the listener to resolve the ambiguity.

If the grammar allows for singleton restrictions in cases like we have discussed in the previous section, then we have to admit this possibility for simpler cases like those with which children typically produce q-spreading responses.

Let's suppose our context contains a set of three boys and four wagons: Boys = $\{B_1, B_2, B_3\}$, Wagons = $\{W_1, W_2, W_3, W_4\}$, as shown in (24). When the child hears the puppet uttering (25) in this context, she has to make a decision about how to restrict quantifier domains. While it is straightforward for the domain of the universal (the child has no reason to consider anything but the entire set of boys shown in the picture because of their visual uniformity), there are at least two equally plausible options for the domain of the indefinite, one of which is restricting it with the salient single wagon:

(24) 

(25) Puppet: Every boy is pulling a wagon.
Child: Can he be talking about this one? 

(26) C=  C'=  or 

Under the surface scope construal, the sentence would have the meaning as in (28). If the child constructs C' to include multiple wagons, then the truth conditions would require that for each boy x there exist a (possibly different) wagon y, such that x pulled y.

(28) $\forall x$ [boy(x) & x $\in$C $\rightarrow$ $\exists y$ [wagon(y) & y $\in$ C' & x pulled y]]

"For every contextually relevant boy x, there is a wagon y in C' and x pulled y."

However, if the domain is narrowed to a singleton set, then the truth conditions for the sentence under the universal wide-scope construal would be effectively the same as under the indefinite wide-scope as in (29); and each boy x would be required to pull the same wagon y.

(29) $\exists y$ [wagon(y) & y $\in$C' & $\forall x$ [boy(x) & x $\in$C $\rightarrow$ x pulled y]]
"There is a wagon y in C' such that every contextually relevant boy x pulled y."

This allows us to explain why children sometimes judge universally quantified sentences like that in (25) false in the context of a picture analogous to the one in (24): unless sufficient linguistic context is given, the salient single wagon in the picture leads the child to construe the expression *a wagon* as a singleton indefinite, the interpretation under which (25) is false in this context.

1.6. Discussion

If we adopt this explanation for children's responses, we are able to explain a large number of empirical findings, some of which were discussed in chapter 1. One observation was that children are not limited to 'spreading error' responses, but rather produce them (in minimal-context tasks) only around 57% of the time (Philip 1995). This is not surprising under this analysis because both a Yes- and a No-response are

grammatically available alternatives. Thus, the 'q-spreading ("No") responses are those given when the child focuses on the perceptually salient single wagon, and interprets the indefinite as singleton, resulting in the "collective" reading, when every boy is required to pull that particular wagon. The adult-like ("Yes") responses are those given when the child chooses the multiple wagons as the domain restriction for the indefinite, hence getting the distributive reading.

This approach also provides a straightforward explanation for the contrast in the rate of q-spreading errors between picture-context and rich-context tasks. Since with the former, the child receives no clues as to what reading the experimenter is expecting, she is left to choose either option. On the other hand, experiments that provide rich linguistic context, as in the Truth Value Judgment task, provide the child with pragmatic clues that are used in natural discourse to interpret and verify a scope-ambiguous sentence, thus limiting the interpretive possibilities and resulting in much lower error rates.

Finally, we can also account for the finding that the one-extra-object condition is harder for children than the multiple-extra-object condition. This is because in the former there is a perceptually salient candidate for the singleton restriction for the indefinite, while in the latter there isn't. Hence, in the former, the child may be biased towards the singleton reading, while in the latter she isn't and would more consistently choose the distributive reading. This can also explain why the extra-object-condition is harder than other conditions, such as those in which the picture contain an alternative agent.[6] This is because according to this proposal, the source of difficulty is the interpretation of the indefinite object and not the universal subject.

---

[6] For example, the sentence *Every boy is pulling a wagon* paired with a picture containing a girl pulling a wagon in addition to boy/wagon pairs.

The question arises why adults would almost never give a No-response in the situations when children do. The difference between adults and children may be due to the differences in their pragmatic and also in their processing flexibility (the former due at least in part to the differences in ToM). I suggest that on-line adults go through the same calculations as children entertaining the same options, but, due to their greater pragmatic flexibility, they are more efficient in selecting the correct readings even with poor context (and/or in rejecting pragmatically implausible readings). Another factor that may be causing the difference between adults and children is in their respective use of the Gricean conversational principle – the Principle of Charity (Grice 1975, Davidson 1984). According to this principle, when given a choice between assigning a reading compatible with the presented situation or the one incompatible with it, adults select the former due to their expectation that the speaker is being truthful. It is possible that for a certain reason children do not observe the Principle of Charity (cf. Hulsey et al 2005). Yet another factor from which the difference between adult and children's behavior may stem is their different degrees of processing flexibility. It allows adults to switch from one reading to another more easily compared with children, who may have difficulty revising their initial parse because they lack computational resources to evaluate multiple options. As a result, they select the first grammatical interpretation they access and stay with it (cf. Lidz 2007). Thus, the difference between adults and children is likely to be due to interplay between various performance factors – both pragmatic and processing, and not in formal semantic knowledge.[7]

[7] Although without additional research it is impossible to separate children's pragmatic inflexibility from their processing inflexibility and to say definitively what is a greater factor in 'q-spreading' – children's inability to cope with sub-optimal pragmatic conditions, their ignorance of the Charity Principle or their inability to revise their initial parse, I believe the totality of the available data

There is independent evidence concerning on-line behavior of adults, which comes from an eye-tracking experiment conducted by Meroni et al. (2001) with English-speaking adults. Their results indicate that adults evaluate alternative domains for the indefinite.[8] This study examined the on-line patterns of fixation duration by adults in response to the extra objects in a picture verification task. Their results indicate that adults not only gazed at the single extra object, which is irrelevant for the calculations of truth conditions of the sentence under the distributive reading, but that the duration of their gaze was significantly longer if the picture contained a single extra object compared with the condition in which there was another object next to it. This result is reminiscent of the Sugisaki and Isobe (2001) child study, in which the number of extra objects in the picture was manipulated. It shows a parallel between adults' on-line and children's off-line behavior. For adults, the gaze duration on the extra object decreases, while for children the error rate decreases if the visual salience of the single extra object is decreased.

2. Issues of Scope

I have proposed looking at the role of the indefinite NP in the so-called q-spreading responses, instead of focusing on the universal quantifier, as has been done previously. As I discussed in the previous chapter, what appears to be 'q-spreading' responses may

---

suggests that children's pragmatic inflexibility does play a role. Thus, since pragmatic manipulations, such as manipulating the visual salience of the extra object or providing enriched linguistic context, trigger changes in children's performance, it suggests that we cannot account for all of children's behavior by saying that the errors are due solely to their selecting one interpretation (at random) and not being able to revise it. Since the experimental data I am reviewing here does not deal with children's on-line behavior, I will leave the role of processing in children's quantification errors for future research (cf. Lidz 2007) and discuss only their pragmatic weakness.

[8] Although the authors themselves do not make this connection, I believe that their results are best explained this way.

simply be errors in specificity, when the child construes the indefinites as singleton in contexts when such interpretation is not intended by the experimenter. As I discussed in section 1.4, 'singleton' indefinites affect the semantic scope relations independently from the syntactic mechanism of relative scope (i.e. the covert movement operation altering the surface c-command relations to determine the relative scope of the scope bearing elements in the structure, or QR). In particular, singleton domains have the effect of neutralizing scope in those structures, which otherwise would have indefinite narrow scope. These indefinites, which would otherwise have the non-specific (also referred to as 'quantificational') interpretation, when combining with a singleton domain have the appearance of being referential and essentially render the sentence scope-neutral (in Schwarzschild's terminology), thus obscuring the surface syntactic scope relations.

Thus, if a listener judges (30) true in a situation in which a number of boys kissed the same girl, we cannot say with confidence that the said listener has the knowledge of "QR" and accepts the inverse scope reading of (30). This is because in this case, one would not need to reverse the scope relations syntactically in order to obtain the truth conditions compatible with the indefinite wide scope reading. One would merely have to narrow the set of girls in the context to one relevant girl. In this situation, the sentence is true under the surface scope, as well as under the inverse scope – what Schwarzschild calls 'scope neutralizing'. Hence, the 'true' judgment cannot be taken as an indication of the syntactic scope reversal. In order to test for that, one needs to use a sentence in which the object quantifier can take wide scope only through some type of covert syntactic movement, such as the example in (31):

(30) Every boy kissed a girl.
(31) A boy kissed every girl.

This state of affairs leads us to another logical conclusion. Namely, it suggests that just as the 'true" judgment for (30) in a situation compatible with inverse scope does not guarantee the availability for the listener of the genuine inverse scope, a 'false' judgment in the same situation may not indicate the unavailability of such interpretation, but stem from the

contextual conditions failing to clearly establish a restriction that characterizes a singleton set.

I believe that this ability of indefinites to neutralize scope through their contextually determined singleton domains may help explain much of the acquisition data not only dealing with 'q-spreading', as discussed previously, but also with children's treatment of scope ambiguities.[9] One point it allows us to make is the importance of dissociating the issues of syntactic scope and issues of contextual domain narrowing, obscuring and interfering with the syntactic scope. Studies that investigate children's knowledge of scope ambiguities must be more precise in asking whether the non-adult-like scope preferences they report are due to children's limitations in applying QR, their non-adult-like pragmatic strategies in processing context when setting contextual parameters, or perhaps their non-adult-like biases for (or against) singleton indefinites. Another, related, point is that the existence of a great variability of results reported in the cross-linguistic literature on scope acquisition is likely to be due to the contextual parameters involved in setting the domains of indefinites, which depend on the pragmatic conditions of each study.

Being able to pinpoint the source of children's interpretive errors by dissociating the syntactic and semantic scope is an important theoretical issue because it would help resolve the question of whether the widely reported scope limitations are a function of children's grammatical knowledge (e.g. is due to their lack of the knowledge of "QR") or the result of applying a pragmatic strategy in responses to some pragmatic condition of the experimental method (as well as a processing strategy). I would like to suggest that what looks like as a semantic or a syntactic limitation in applying "QR" may in fact be a manifestation of an incorrect assumption with respect to the size of the quantifier domains.

The question of whether children under the age of 5 possess the grammatical knowledge necessary for deriving scope ambiguities was the central question raised by

[9] Again, this analysis is compatible with other theories of indefinites, such the choice function approach; however, I will use the 'contextualist' theory of indefinites by Schwarzschild because it lends itself very well to explaining children's data I will be discussing here.

Musolino in his 1998 dissertation. His investigation was focused on sentences with negation. The acquisition puzzle that he worked to resolve was how children get to know the difference with respect to scope ambiguities between the following sentences (the adult scope possibilities are given in parenthesis):

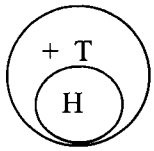
(i) Cookie Monster didn't eat two cookies. (neg<two; two<neg.)
(ii) Cookie Monster didn't eat some cookies (neg<some*; some<neg)
(iii) Cookie Monster didn't eat every cookie. (neg<every; every<neg*)
(iv) Every monster didn't eat a cookie. (every<neg; neg<every)
(v) Some monster didn't eat a cookie. (some < neg; neg<some*)

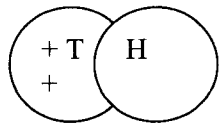
These minimally different sentences have distinct interpretive options for English-speaking adults, where (i) allows both the wide and narrow scope of the indefinite NP, (ii), on the other hand, only allows the indefinite object to have wide scope; (iii) is limited to negation taking wide scope, while in (iv) both scope readings are permitted; finally in (v) negation is blocked from taking wide scope.[10] Given such an inconsistent paradigm, and since, as Musolino observes, it is unlikely that children get explicit instructions as to what meaning is permitted and what meaning is banned for each sentence, it is quite mysterious (aside from appealing to "magic powers") that children manage to come to possess all and only correct interpretive options. The answer to the puzzle proposed by Musolino is that in this task, children are helped by a learning constraint that limits their initial interpretation to the readings derived from the surface scope relations. Children are unable to hypothesize the full adult system at once; and their limitation is not a handicap, but an advantage. Their initial hypotheses are limited in such a way that ensures that any errors in the initial hypothesis entertained by the child can be corrected on the basis of positive evidence alone.

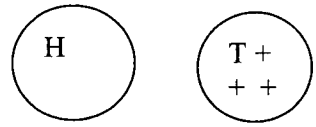
Thus, if children initially allow only those scope readings that are isomorphic to the surface c-command relations in each sentence in the paradigm above, they would have correct interpretive options for (i), (iii), (iv) and (v) (although incomplete for i and iv), and

[10] For some adult English speakers, the sentence in (iii) can have the inverse scope, and the one in (ii) can have the surface scope if the sentence is an emphatic denial. In the "out-of-the-blue" context with a neutral intonation, both sentences are limited to the surface and inverse scope reading respectively.

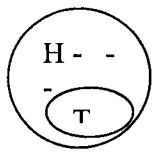
an incorrect one for (ii). However, positive evidence alone would be sufficient for them to learn that the hypotheses they are entertaining for (i) and (iv) are incomplete and those for (ii) incorrect. This theory is based on the idea that if children started out without the constraint limiting them to isomorphic readings, they would face the so-called "superset problem". This problem stems form the observation that in the process of acquisition children do not receive reliable and consistent information about unacceptability of a given sentence (Brown and Hanlon 1970, Marcus 1993). In this situation, if a hypothesized grammar (H) generates a subset of the expressions in relation to those generated by a target grammar (T) as is illustrated in diagram (a), positive evidence will suffice for the child to be able to get rid of erroneous hypotheses. The same holds if H partially intersects with T (b), or does not intersect with it (c):

(a) 

(b) 

(c) 

However, if the hypothesis is a superset of the target, as in (d), a learnability problem arises because every structure of the target is also a structure of the hypothesis and in the absence of negative evidence, nothing in the input would tell the learner that 'excess' structures are not part of the target (Lasnik 1989).

(d) 

The logic of the theory of learnability, then, would require the child to be limited to the grammar that either generates a subset of sentences allowed by adult grammar or a non-intersecting (or partially intersecting set), but never a superset. Based on this theory, Musolino hypothesized that being limited to isomorphism would give the child the required subset of readings with respect to (i) and (iv), and non-intersecting set of readings with respect to (ii). The experimental evidence collected by Musolino 1998 and several

subsequent studies supported his Isomorphism hypothesis. However, a number of other studies, reviewed in the next section, produced results inconsistent with the Observation of Isomorphism and showed that children allow inverse scope interpretation for sentences with negation (Lidz et al. 2004, Miller and Schmitt, 2005 Lidz and Musolino 2002, Gualmini 2004, Hulsey et al. 2005). Hulsey et al. offers a pragmatic explanation for why previous studies were unable to elicit inverse scope readings reviewed in the next section, pointing out an experimental flaw responsible for the isomorphism-limited responses. Lidz et al. contribute an additional processing factor that may lead to the 'isomorphism effect' in children. However, the initial puzzle raised by Musolino remains. If scope ambiguities are available to children from early on, what ensures that if they hypothesize ambiguities for sentences that do not allow them in adult grammar, that they would they be able to successfully unlearn incorrect interpretations? The sentences for which this problem can potentially arise are (ii), (iii), and (v) repeated below:

(ii) Cookie Monster didn't eat some cookies (neg<some*; some<neg)
(iii) Cookie Monster didn't eat every cookie. (neg<every; every<neg*)
(v) Some monster didn't eat cookies. (some < neg; neg<some*)

In the previous discussion, I have addressed the role of 'singleton indefinites' in creating an illusion of a semantic anomaly in children's interpretation of the universally quantified sentences. Now, I would like to have a brief discussion of the latter – the role of quantifier domain restrictions in creating the illusion of children's non-adult-like scope preferences, and then formulate some predictions testing which will be the goal of the experiments reported in the next chapter. I will argue that if we factor in the ability of indefinites to have Schwarzschild-style domain-narrowed readings, we can achieve consistency in this paradigm, simplify the acquisition task facing the child, so that we no longer face the superset problem by allowing the child to have access to scope ambiguities from early on and acquiring the all and only adult-like scope readings. The unavailable readings would have overt structural information to rule them out; the structural information that children demonstrably are aware of. On the other hand, this approach will explain why

certain experimental tasks are hard for children and help explain away some of the seemingly semantically anomalous responses reported in the literature on children's knowledge of scope, as we did with quantifier-spreading.

2.1. Do children lack some scope readings?

The literature on the acquisition of scope roughly speaking follows two different approaches:

- Those studies that look at scopal ambiguities as an issue at the syntax/semantics interface; i.e. a function of children's knowledge of "QR" and the language specific conditions permitting or limiting its application (e.g. Musolino 1998, Musolino, Crain, and Thornton 2000, Lidz and Musolino 2002, Gualmini 2002);
- Those studies that look at children's ability to access both 'referential' and 'non-referential' indefinites as an issue of lexical acquisition independent from the issue of syntactic scope (based on the lexical ambiguity view of indefinites by Fodor and Sag (1982) or other compatible with the ambiguity view accounts (e.g. Kramer 1998, 1999, deHoop and Kramer 2006, Miller and Schmitt 2004).

Interestingly, the conclusions about children's scopal preferences drawn in studies from the two schools of thought are often contradictory. Thus, it has been frequently reported in the "QR literature" that children exhibit isomorphism in their scopal preferences – a limitation that causes them to accept the surface scope interpretation for sentences that are scopally ambiguous or even for those sentences in which adults strongly prefer the inverse scope reading. Thus, Lidz and Musolino (2002) tested preschool-age English- and Kannada-speaking children (the mean age for English was 4;4 and for Kannada 4;5) on their interpretation of ambiguous sentences involving quantified noun phrases and negation, as in (33).

(33) a. Donald didn't find two guys.
b. Anoop eradu kaaru toley-al-illa.
Anoop two car drive-inf-neg
"Anoop didn't drive two cars."

Both the English and Kannada sentences are ambiguous between the surface scope interpretation, in which the indefinite is interpreted under the scope of negation, and the

inverse scope reading, in which the indefinite takes wide scope. The study confirmed the existence of the scopal ambiguity by showing that both English- and Kannada-speaking adults readily accessed either interpretation. However, both English- and Kannada-speaking children responded in a way consistent with the "isomorphism hypothesis" - accepting the sentence in a situation consistent with its surface scope reading (e.g. the agent did not find any guys in 33a or drive any cars in 33b), but rejecting it in a situation consistent with the inverse scope (when the agent found some guys or drove two cars, but failed to find or drive two of them). Similar isomorphic responses were found in Lidz and Musolino 2004 with respect to the sentences in which the surface scope configuration contained negation within the c-commanding domain of the indefinite, as in (34):

(34) Two butterflies did not go to the city.

(35) eraDu chitte paTNa-kke hoog-al-illa
two butterflies city-DAT go-INF-NEG
'Two butterflies did not go to the city.'

(36) $\exists_{2x}$ [butterfly(x) &¬[x went to the city]] **Surface scope interpretation**
(37) ¬$\exists_{2x}$ [butterfly(x) & x went to the city] **Inverse scope interpretation**

When sentences like (34) and (35) were presented to children in the context of a story consistent with the surface scope reading (e.g. a story in which two butterflies decided to go to the city, while two others decided to stay in the forest), the sentence was accepted as true at the rate of 87.5%. However, in the inverse scope context, (e.g. a story in which one butterfly went to the city and one butterfly stayed in the forest), the rate of acceptances dropped to the mere 20%. Similarly, 'isomorphic' responses were reported by Musolino 1998, Musolino, Crain, and Thornton 2000, Musolino and Gualmini 2004 among others.

In short, the 'isomorphism' literature sought to demonstrate that children's limitations lie in their ability to access those readings that require a reversal of the surface scope relations, irrespective of the specific syntactic or semantic structure under consideration, and hence, are not a result of their bias towards either wide or narrow scope for indefinites or other quantifiers per se. On the other hand, the literature in the lexical ambiguity tradition frequently reported that the non-adult-like responses with respect to semantic scope of sentences containing indefinites reveal children's biases towards a certain interpretation of indefinites, even when such interpretation is non-isomorphic, and moreover, if such non-isomorphic interpretation is dispreferred (or even disallowed in the adult grammar), thus directly contradicting the Observation of Isomorphism.

Lee (1986) investigated children's understanding of sentences with single quantifiers such as *yi-ge ('a')*, *dou ('all')*, *quan ('all/entire')* in Chinese, *every* and *all* in English, as well as the interaction of two quantifiers using picture identification and/or act-out tasks. Lee observed that for a sentence like (38), whose preferred adult interpretation is consistent with the surface scope reading and is given in (39), Chinese-speaking children as old as 8 interpret the indefinite *yige dangao* ('a cake') as referring to a specific entity. Such rejecting the adult-like isomorphic interpretation and instead preferring the non-adult-like inverse scope construal is contrary to Musolino's observation of isomorphism.

(38) Meige xiaohai dou zai chi yige dangao
every-CL child all ASP eat one-CL cake
'Every child is eating a cake'

(39) For all x=child, there is a y=cake such that x is eating a cake.

On the basis of such findings, Lee concluded that Chinese-speaking children "probably interpreted QNPs as if they are inherently referential" (p.190).

Wide-scope indefinite bias was also reported for Dutch-speaking children. The status of Dutch-speaking children's scope preferences, however, became a subject of some

controversy. Thus, Bergsma-Klein (1996) reported that for sentences like (40), which for adults have unambiguously isomorphic scope interpretation, Dutch-speaking children are adult-like in correctly assigning a referential ('wide scope') reading to indefinite subjects.

(40) Een meisje gleed twee keer uit.
A girl slipped two time out-PARTICLE
"A girl slipped twice."

However, Termeer (2002) reported results that complicated the picture of Dutch-speaking children scope preferences. Termeer looked at the interpretation of Dutch sentences with the so-called embedded subjects, as illustrated in (41), for which adults again strongly prefer the isomorphic interpretation - in this case a 'non -referential' indefinite reading (e.g. the reading in which there were two different boys who went down the slide). She found that for this structure, children again exhibited a strong preference for the non-isomorphic 'referential' indefinites – a reading in which one boy goes down the slide twice. In Termeer's sample, only 32% of the children (ages 8;7 to 10;4) accepted the adult-like narrow scope reading, with the majority interpreting the sentence as referring to one boy going down the slide twice.

(41) Er ging twee keer een jongen van de glijbaan af.
there went two time a boy of the slide off
"Twice, there went a boy down the slide."

Thus, Dutch child data presents a puzzle: in a language that obeys a constraint that limits scope ambiguities and disallows non-isomorphic readings, children, who had been claimed to be developmentally limited to non-isomorphic readings, seem to disobey the constraint, which should be available to them universally. However, if we look more closely at adult Dutch, we would notice that the nature of the constraint limiting adults to surface scope readings is not well understood and may not be of a grammatical, but a pragmatic nature (Philip 2006). Thus, it's long been known that in Dutch the constraint limiting subject indefinites to the specific interpretation is not absolute, and the indefinite subject of any

transitive predicate can be nonspecific (Reuland 1988). In particular, contextual factors can cancel the constraint that causes indefinite subjects to be interpreted 'referentially', suggesting that the constraint itself is not a syntactic limitation on the application of QR. For example in (42) the subject need not be specific (true with more than 1 dancing dwarf):

(42) Een kabouter heeft op iedere tafel gedansd
a dwarf has on every table danced.
'A dwarf has danced on each table.'

As van den Wijngaerd (1992) has noted, a Dutch sentence like (43) does allow a wide scope reading of the universally quantified object, just like its English counterpart.

(43) Een waakhond bewaakte ieder huisje.
'A watchdog was guarding each house'

It is outside the goals of the present study to investigate the precise mechanism responsible for scope interactions in Dutch. However, the examples like (42) and (43) suggest that firstly, both 'referential' and non-referential' readings are possible for indefinite subjects; secondly, which of the two available readings is selected by adults is context-dependent; and thirdly, in the absence of context the 'referential' reading is selected by default. A theory of indefinites that handles the 'referential /non-referential' ambiguity as a function of the size of domain restriction, which is contextually determined, would be able to handle both the adult and child Dutch data in the most parsimonious and explanatory way. Furthermore, it would be able to handle those experimental findings that show Dutch children to be limited to non-isomorphic readings, but in ways inconsistent with the results reported by both Termeer and Bergsma-Klein.

Thus, some studies revealed that Dutch-speaking children not only rejected adult-like isomorphic scope interpretation, contrary to the isomorphism hypothesis, but that their non-adult-like non-isomorphic interpretation resulted from children's preference for the 'non-referential' or narrow scope indefinites, in direct contradiction to the studies like Lee 1986,

Bergsma-Klein (1996), and Termeer 2002, among others. Thus, Krämer 1999 studied children's interpretation of indefinite subjects in sentences like (44):

(44) Een meisje is niet aan het dansen.
a girl is not PROGRESSIVE dance
'A girl is not dancing"

As in example (40), this sentence receives an obligatory wide-scope indefinite interpretation in adult Dutch. This means that when adult subjects were shown a picture with one girl dancing and one girl not dancing (e.g. swinging on a swing), 100% of them accepted the sentence as true. In contrast, 48% of the 4-5 year-olds rejected (44) with another 26% exhibiting a mixed pattern rejecting some and accepting some sentences in this experimental condition. Similarly, Krämer 2000 observes that for Dutch sentences like (45), which adults interpret as indefinite wide-scope, children assign a non-specific reading to the subject. Being asked to act out (45), children rolled two different marbles, whereas adults only roll one marble twice.

(45) Een knikker mag twee keer rollen.
'A marble may roll twice.'

Krämer 2000 showed similar results with indefinite objects. Dutch object noun phrases can have either a wide-scope or narrow-scope reading depending on the surface position of the indefinite. If the indefinite occurs lower than negation, it has an obligatory narrow-scope reading in the adult grammar (as in (46). If it occurs higher than negation, it has an obligatory wide-scope reading (as in (47).

(46) De jongen heeft geen vis gevangen. (neg > a; *a > neg)
The boy has not+a fish caught
"The boy didn't catch a fish."

(47) De jongen heeft een vis niet gevangen. (*neg > a; a > neg)
The boy has a fish not caught
"The boy didn't catch a fish."

The results of Krämer's study showed that Dutch children between the ages of 4 - 7 years old (mean: 5;8) did not distinguish between these two sentence types, and instead,

assigned narrow-scope readings to both scrambled and unscrambled object noun phrases. Krämer attributed this response to the child's difficulty with "discourse integration", a difficulty, which according to Krämer lies in the integration of the specific referent in the accommodation site of the common ground.

Philip and Botschuijver (2004) further investigated the Non-Integration hypothesis by testing it with English-speaking children. They conducted two experiments - an act-out and a TVJ task –testing sentences analogous to those used by Krämer 2000. They compared the rates of the 'referential' interpretation that English-speaking children and adults assign indefinite subjects in sentences in which they are preceded with a scope-taking adverb like 'twice'. Even though in English they allow both 'referential' and 'non-referential' readings, P&B reported that in the act-out task, both adults and children strongly preferred to interpret the indefinite subject as 'referential', casting doubt on the Non-

Integration Hypothesis. However, they got a contrasting result in the truth-value judgment task. While adults continued to favor the referential reading, accepting sentences like (45) as true in the situation in which one marble rolled twice 93% of the time, children accepted the sentence in this context remarkably less often than the adults did. The lowest acceptance rate was reported for the 6-year-old group at only 43%, compared with this group's acceptance rate of 73% for the nonspecific context, in which two different marbles rolled once each. P&B interpreted these findings as support for the Non-Integration hypothesis.

Children's bias for "non-referential" indefinites was implicated in 'q-spreading'. Philip 2004 reported that his Dutch-speaking children often judged sentences with indefinite subjects analogous to (48) false of contexts like (49), in which adults uniformly judge the sentence true, and true of contexts like (50), in which adults reject the sentence as false (for adult intuitions Philip relies on Ruys 2001). Philip interpreted his results as evidence that Dutch-speaking children interpreted the indefinite as 'non-referential' and hence rejected it in the context true of the surface scope.

(48) An elephant is carrying every dog.
(49) 1 elephant carrying 3 dogs + 2 extra elephants.
(50) 3 elephants each carrying a different dog.

Similar claims have been made for Chinese, a language that has a constraint similar to that found in Dutch, requiring isomorphic ('referential') interpretation for unscrambled subject indefinites. Some studies claimed reported that children disobey this constraint and permit 'non-referential' indefinites (Lee 1991, Chien 1994, among others). As we see, the results of cross-linguistic research on children scope preferences contains some contradictory results. Some core findings in this area are summarized below:

Table 1. ***Children's scope preferences***:

I. ***Evidence for Isomorphic Scope Preference:*** **children reject the inverse scope *(wide-scope indefinites) in neg. sentences (*** * indicates results significantly different from those obtained form adults):

English:
Musolino, 1998; Lidz & Musolino, 2002 **(cf. also Musolino and Gualmini 2004)**
The girl didn't find some/two frogs
some>neg: 35%* (mean age: 5;1)
two>neg: 33%* (mean age: 4;4)

Kannada: ***Lidz& Musolino, 2002***
Anoop eradu kaaru toley-al-illa
"Anoop didn't wash two cars"
two>neg: 23%* (mean age: 4;5)

Dutch: ***Kramer, 2000***
De jongen heeft geen vis gevangan
the boy has not+a fish caught
"They boy did not catch a fish"
a>neg: 16% (mean age: 5;8)

II. Evidence for Non-isomorphic Scope Preference (wide-scope indefinites):

Dutch: Termeer 2002
Er ging twee keer een jongen van de glijbaan af.
there went two time a boy of the slide off
"Twice, there went a boy down the slide."

a > twice 68% * (age 8;7 to 10;4)

Chinese: Lee 1986
Meige xiaohai dou zai chi yige dangao
every-CL child all ASP eat one-CL cake
'Every child is eating a cake'
every > a 62%* (cf with the adults, who accepted this interpretation 95% of the time.)

Chinese: Sue 2001
Milaoshu meiyou qi yi-zhi gou.
Mickey Mouse didn't ride a-CL dog
"Mickey Mouse didn't ride a dog."
a > neg 77% (mean: 5;0) (cf. with the adults, who accepted this reading at the rate of 64% with no statistical significance.

III. Evidence for Non-Isomorphic Scope Preferences (Narrow Scope Indefinites):

Dutch: Kramer 1998
Een meisje is niet aan het dansen
a girl is not PROGRESSIVE dance
'A girl is not dancing"
a > neg 26% (age - 4;0 to 5;6)

Kramer 2000:
De jongen heeft een vis niet gevangen. (*neg > a; a > neg)
The boy has a fish not caught
"The boy didn't catch a fish."
a > neg

English: Philip and Botschuijver 2004
A marble may roll twice.
A > twice 57% (average across all age groups)
Twice > a 72%.

In contrast to this group of studies reporting non-adult-like scope preference, albeit disagreeing as to what non-adult-like property determines children's scope preferences, there are studies that demonstrate that with certain changes in the experimental design, children become adult-like in their acceptance of the readings they were claimed to reject. Thus, in the experiment conducted by Miller and Schmitt (2004), English children were tested for the availability of the wide scope of indefinite objects. In their experiment,

Miller and Schmitt modified the background stories in such a way that made one member of the object set highly salient. They specified the object set (presented a clearly delineated set of individuals from which the relevant one had to be selected), set up the stories in such a way that the protagonist was explicitly required to carry out an action on all of the objects in that set, and finally highlighted the "leftover" object that did not undergo the action by pointing it out right before presenting the target sentence.

For example, in one story, the protagonist was a girl named Kelly, who was told by her mother to paint a basket of eggs for Easter. The girl painted all of the eggs but one. Before presenting the target sentence "Kelly did not paint an egg." the experimenter made the one unpainted egg salient by saying something like the following: "Now Kelly is going outside to play, but look! What about this egg?" (points to the unpainted egg). In this study, English-speaking children (N=33, ages 3;10-5;8) accepted the wide scope indefinite 91.6% of the time, while adults accepted these sentences 97.9% of the time (the difference was not significant). This finding contradicts both Musolino-like results and Kramer's claim about children lacking the 'referential' indefinites at the initial stage of acquisition.

Another study that showed a dramatic improvement of children's performance with a contextual manipulation was Musolino and Lidz (2003). In their study, they tested whether children, who were previously reported to lack the inverse-scope interpretation of sentences such as 'Every horse didn't jump over the fence', could be influenced to accept this reading at a higher rate. They found that the acceptance rate of the inverse scope was greatly facilitated in the presence of a preceding affirmative clause, as in 'Every horse jumped over the log but/and every horse didn't jump over the fence'.

Studies like Miller & Schmitt 2004 and Musolino & Lidz 2003 strongly suggest that whatever interpretive biases may be revealed under certain experimental conditions

do not reflect limited grammatical options since manipulating experimental conditions demonstrably changes the experimental outcome.

To summarize the discussion so far, children have been shown to respond to scopally ambiguous sentences in truth value judgment tests in a way that suggests a preference for isomorphic scope. On the other hand, other studies reported a bias for 'referential' indefinites, when the indefinite is assigned 'wide scope' regardless of its surface syntactic position even if adults dispreferred this reading in a given construction. At the same time, still other studies showed an opposite tendency of assigning narrow scope to indefinites, where adults would either permit or require the wide scope interpretation. The discrepancies in the reported children's scope preference are perplexing and neither the lexical ambiguity nor the "QR" approach provides an explanation as to what component missing from child's grammar may be responsible for such a broad range of scope-related errors. However, the "singleton indefinite' approach, which links semantic scope of indefinites with contextual domain restrictions and obtains wide scope-like indefinites without QR, makes the variability of the reported results across various studies less puzzling and allows us to easily account for each type of error as a function of the experimental conditions under which it has been elicited.

2.2 How can 'singleton indefinites' help?

The 'singleton indefinite' account reveals a problem for studies aimed at testing children's or adults' ability to access inverse scope. It suggests that certain types of quantified sentences, namely sentences in which the inverse scope would require an indefinite to

undergo "QR"[11], are unsuitable for testing the existence of a genuine inverse scope interpretation in either adults or children. In such sentences genuine inverse scope would be indistinguishable from scope-neutral singleton indefinite readings discussed by Schwarzschild and reviewed in section 1.4 because of the option to derive what appears to be syntactically inverse scope from the surface scope by simply narrowing the domain restriction of the indefinite. This means that any study that shows either a lack or a preference for a certain scope reading by using such a sentence cannot be used reliably to answer the question about the available scope readings. What they likely to measure is a strategy used by children or adults when selecting a reading of a scope-ambiguous sentence in a given experimental situation; e.g. how children differ from adults in using visual context or overcome limitations in linguistic context to resolve scope ambiguities.

Let's look at some of the puzzles arising from the studies on acquisition scope. First, let's look at the 'isomorphism' studies that (i) show that children are biased towards isomorphic scope interpretation even for those structures for which adults strongly prefer the inverse scope reading (e.g. Musolino 1998); and (ii) such seemingly unavailable reading can be greatly facilitated by certain experimental manipulations (Lidz and Musolino 2003 etc. cf. also Hulsey et al. 2005)[12]. One of the major "isomorphism" studies was Musolino 1998, who investigated children's knowledge of and the acquisition path with respect to the following paradigm repeated below from section 3.2 (the permissible adult interpretations shown in parenthesis):

(i) Elmo didn't catch two butterflies. (neg< two; two< neg)
(ii) Elmo didn't find some butterflies (neg<some*; some< neg)
(iii) Elmo didn't catch every butterfly (neg< every; every<neg*)
(iv) Some horse won't jump over the fence. (some <neg; neg<some*)
(v) Every horse didn't jump over the fence. (every < neg; neg < every)

---

[11] I will use the term "QR" in a theory-neutral sense to refer to a covert syntactic mechanism of scope reversal without committing to any particular implementation of this mechanism.

[12] Their studies usually used sentences with a plural indefinite or a numeral quantifier, which I believe can be handled with the same analysis of specific/non-specific ambiguity as indefinite singular adjusted for cardinality of the domain restriction. Thus, the 'wide-scope' reading of the sentence equivalent to Fodor and Sag's example (e.g. "If two relatives of mine die in a fire, I will inherit a house" is derived from the surface scope by narrowing the domain of the indefinite to two relevant individuals. I will use singular indefinites for consistency.

Children's responses followed the following pattern:

a) adult-like interpretation for (iii) and (iv) (highlighted in green);
b) accepting a subset of the adult readings, namely the isomorphic one, in i and v (highlighted in red);
c) rejecting the adult reading, while accepting the non-adult-like reading in ii (highlighted in purple);

The two immediate questions that arise with respect to this pattern of children's responses have to do with (i) the difference between the sentences that leads to the difference in children's adult-like responses – why do children behave like adults with respect to some of them, while differ from adults with respect to others; and (ii) what mechanism would allow children to eventually converge on the correct adult interpretations. The latter question is important with respect to sentences like (ii), for which children were reported to have the reading ruled out for adults. As Musolino observed, since the reading unavailable for (ii) exists for all other sentence types, and since the information about ungrammaticality of sentences is not available to children, it becomes a puzzle to explain how children eventually learn that the surface scope for (ii) is unavailable. The proposed solution, as reviewed earlier, was to say that children start out lacking scope ambiguities with only isomorphic scope readings available to them. In this situation, sentence (ii) does not pose the superset problem. Since the children's interpretation and the adult interpretation represent non-intersecting sets, positive information alone would suffice to tell the child to revise their grammar. This elegant solution, however, has an empirical problem since numerous studies not only show that children's responses may be biased to certain non-adult-like non-isomorphic readings (e.g. Lee 1986, Termeer 2002), but also that given certain manipulations of the experimental conditions, children respond consistently with scope ambiguities (Miller and Schmitt, Hulsey et al.). If children know from early on that at least in some cases scope relations in a sentence may be reversed, then sentence (ii) should still present the 'superset problem', which Musolino 1998 was trying to solve.

Let's look at the paradigm again and try to see whether we need not worry about children reaching adulthood with the erroneous idea that (ii) is ambiguous between the

surface and the inverse scope readings. The immediate observation from looking at the pattern of children's response is that apart form the problematic (ii), the sentences that are unambiguous for adults are also unambiguous for children, while the scope-ambiguous sentences present some difficulty and children frequently fail to access the inverse scope readings. Therefore, the question that needs to be answered with respect to these sentences is how children know that (iii) and (iv) are not ambiguous and why they failed to access the inverse scope readings with respect to (i) and (v).

First, let's address the latter question and explain why children failed to accept the non-isomorphic readings in (i) and (v). A very convincing explanation for this question was proposed by Hulsey et al 2005, who attributed children's failure to accept legitimate non-isomorphic readings for sentences like (i) to a certain methodological flaw. In experimental studies like Musolino 1998 and those analogous to it, the authors rely on the assumption that in order to check whether a child allows a particular interpretation, all that's required is to present the ambiguous sentence in a situation that makes the targeted interpretation (e.g. its inverse scope reading) true while simultaneously making other possible readings false. This assumption is based on the empirical observations about adults, who had been reported to meet this expectation, known in the literature as the "Principle of Charity" (Musolino and Lidz 2003). When children fail to accept the target sentence under these conditions, the response had been interpreted to indicate a lack of the reading in the child's grammar. It is, however, possible that the child rejected the target reading not because it is disallowed by his grammar, but because he was answering not according to the Principle of Charity, either because he hasn't acquired it yet, or because another, more robust pragmatic factor overrode it.

Hulsey et al. 2005 suggest one such pragmatic factor interfering with the Principle of Charity. They provided an explanation for Musolino's isomorphism findings by looking at a condition that comes into play when we choose an alternative while interpreting ambiguous sentences and which may overwride the Principle of Charity. This condition

requires that the interpretation selected, whether true or false, be a good answer to a certain Question under Discussion, a question made salient by the preceding context. According to this hypothesis, a story told in a Truth Value Judgment Task always raises an implicit question, so that the target sentence presented in conjunction with the story will be interpreted as an answer to that particular question, which may or may not be the intention of the experimenter. The reason why children often fail to give adult-like truth value judgments is that the experimenter neglects to control for the question made salient by the story-context. When both interpretations of a target sentence happen to address the Question under Discussion, then children, like adults, will adhere to the Principle of Charity and choose the interpretation that makes the target sentence true. However, if only one interpretation addresses the Question under Discussion, children will evaluate the truth-value of the target sentence construed in the appropriate way, regardless of whether that interpretation makes the target sentence true or false. Importantly, the choice of any given interpretation is not dictated by the relationship between surface c-command and scope. According to this hypothesis, the contexts in which children have been found to give isomorphic responses fail to establish an appropriate Question under Discussion (the one focusing on the inverse scope reading).

This explanation can account for the pattern of responses found by Musolino 1998 we are trying to explain. Children were limited to the isomorphic reading with respect to (i) and (v) because the context in which they were presented fail with respect to the relevant Question under Discussion. In order to elicit the non-isomorphic 'neg < every' reading in (v), the story context has to address the question "did **every** horse jump over the fence?". Appropriate answers to this question are 'yes, every horse jumped over the fence" and 'no, not every horse jumped over the fence". In this situation, (v) would be most likely interpreted under the inverse scope, which is synonymous to the latter alternative. On the other hand, if the story doesn't presuppose this question or any particular question, the

default question the child would consider is "what did the horses do?", in which case the surface scope would be the more appropriate answer.

Similarly, the non-isomorphic '*two* < *neg*' reading for (i) would require a story that leads one to answer the question "How many butterflies did Elmo *not* catch?". Simply telling a story in which Elmo catches some butterflies and does not catch others fails to do that. Instead, such a story raises the question "did Elmo catch any butterflies?" or "how many butterflies did Elmo catch?". An appropriate answer to this question is either "Elmo caught an x-number of butterflies" or "Elmo did not catch *any* butterflies". Of the two available readings, the surface scope reading of (i) works as the more appropriate answer to this question (since it is a denial that Elmo caught an x-number of butterflies) and hence becomes the favored response elicited form children.

An example of a study that successfully addresses the problem is Gualmini (2003a), who deliberately set up the story-contexts in such a way that they brought up an appropriate Question under Discussion necessary for the inverse scope construal of the target sentences. Thus, in one story the children were told that Grover ordered four pizzas from the Troll. They were also told that he was supposed to deliver all four of them, but was driving too fast and lost two of the pizzas on the way. In this context, the target sentence uttered by the puppet was addressing an implicit question of whether the Troll managed to deliver all of the pizzas. The two possible outcomes are (52) and (53), which correspond to the Hamblin denotation of question in (51) (Hamblin 1973).

(51) Will the Troll deliver all of the pizzas?
(52) Yes, the Troll will deliver all of the pizzas.
(53) No, the Troll will fail to deliver some of the pizzas.

The target sentence uttered by the puppet is given in (54):

(54) The Troll didn't deliver some pizzas.

Here, even though (54) is ambiguous between its surface scope readings (false) and the inverse scope readings (true), only the inverse scope reading addresses the Question under

Discussion. The two readings can be paraphrased as (55a,b):

(55) a. It is not the case that the Troll delivered some pizzas.
= The Troll didn't deliver any pizzas.
b. There are some pizzas that the Troll didn't deliver.

Since, the context explicitly makes (51) the Question under Discussion, the children interpreted the target sentence as an answer to this question and chose the reading consistent with being an appropriate answer. Consequently, Gualmini's children were successful in accepting the non-isomorphic interpretation of (54), in contrast to Musolino's children, who rejected equivalent sentences in an equivalent context.

In light of these new findings, we are led to conclude that children's weakness exposed by the 'isomorphism' studies is not in the area of their grammatical competence, but is a limitation in their ability to figure out communicative intentions of the speaker in situations in which these intentions are left vague. In an experimental situation, when choosing among the available readings of an ambiguous sentence, adults are able to figure out that the logic of the experiment requires them to interpret the target sentence in a way that 'disagrees' with the implicit Question under Discussion. In other words, they understand that passing the test requires them to suspend the Question under Discussion and search for another, less obvious interpretation. Being much more skilled in test taking, adults know that in any test, certain elements are present in the testing situation simply to distract, and therefore require a savvy test-taker to ignore them. Children, on the other hand, are "led down the garden path" by the question implicitly raised by the experimental situation. This can remedied by manipulating the context in such a way that it makes the target reading more salient (e.g. by explicitly stating the appropriate Question under Discussion, as discussed by Hulsey et al. or by adjoining the negated sentence to a clause affirming a proposition contrastive to what the inverse scope of the negated sentence denies, (e.g. 'Every horse jumped over the log' facilitated children's acceptance of the inverse scope of "every horse didn't jump over the fence"), as in Lidz and Musolino 2003).

This approach successfully explains why children responded the way they did to sentences (i) and (v) in Musolino's study (and presumably other studies reporting children to be limited to one of the readings available to adults). Another question that we need to answer is whether children's isomorphic responses with sentences for which adults are also limited to the isomorphic interpretation (iii) and (iv) are a result of their knowledge that these sentences do not permit inverse scope, or whether they simply respond isomorphically due to the experimental conditions being consistent with such responses, analogous to the cases discussed above. If these sentences under some conditions exhibit scope ambiguities for children, the superset problem would arise.

In (iv), the absent reading is the one that would land the negation in a position from which it would c-command a positive Polarity Item "some", and hence "anti-license" it. Musolino did not test directly whether children would accept (iv) in context compatible with their inverse scope. Hence, we cannot conclusively say whether children lack this reading due their knowledge of the rules governing the distribution of PPIs or whether they may allow this reading under some conditions. In (iii), which is also unambiguous for adults, nothing should block "QR" and the sentence should be expected to have both readings. The difference between this sentence and the one in (iv) is that in the former, the negation is lower than the quantifier in the surface structure, while in the latter it is higher. Therefore, for the former it is negation that would have to raise to a position above the quantifier, while in the latter sentence it is the quantifier that would have to undergo QR to have its scope relations reversed. The latter option for one reason or another is blocked. The reason why the inverse scope for this sentence is blocked is not of direct relevance to the question at hand. Therefore, I will simply adopt this as a descriptive generalization that according to this data, in a situation in which a quantifier interacts with negation, only negation has an option of raising above the quantifier to reverse scope relations, while the quantifier is barred from doing so. This generalization would account for the ambiguity of sentences like (v), as well as for a lack of ambiguity in (iii): in (v) the negation can be interpreted both below and

above the quantifier; in (iii) the quantifier must remain below negation. In (iv), the negation raising above the quantifier should be possible, but is blocked for the reason we have discussed above (anti-licensing of a PPI).

The two remaining cases are (i) and (ii). In (i) children gave isomorphic responses, while both the surface and the inverse scope readings are available for adults. This sentence is only minimally different from the unambiguous (iii). I would like to suggest that the reason why the inverse scope reading is available in (i) is that the indefinite here can be interpreted as a Schwarzschild-style scope-neutral specific indefinite, and not a genuine inverse scope induced by "QR". The difference between (i) and (iii) is in the availability of the Schwarzschild-style specific indefinite in the former but not the latter, explaining why (i) is ambiguous: it has the surface scope with a regular domain (the one that includes all contextually relevant individuals) and a reading in which the domain is narrowed:

(56) $\neg\ \exists_{2x}$ [butterfly(x) & x is in C & Elmo caught x]
(57) C= {x: x is a butterfly} or C={B_1, B_2}

Failure to elicit the "inverse scope" is not surprising and is due to the experimental flaw discussed by Hulsey et al. The remaining case is (ii), which remains mysterious.

In this case, adults do not simply disfavor the surface scope for (ii), but reject it, while children, on the other hand, prefer the reading unavailable to adults. This finding is the primary one that raised the question of how children would avoid the superset problem if scopal ambiguities are available to them from early on and provided the motivation for positing the isomorphism stage, as a way to avoid it.

The two logically possible readings for (ii) are given below. The reading in (a) is reported to be unavailable to English adults, while the reading in (b) is available. Judgments are reversed for children.

(ii) Elmo didn't catch some butterflies.
(a) Elmo didn't catch any butterflies.
(b) Elmo caught some butterflies, but he failed to catch some of the butterflies.

One clue for why this sentence is harder for children despite its lack of ambiguity is

that adults' judgments for sentences like (ii) are quite 'fuzzy', particularly if the sentence is given in an "out of the blue" context. Adults tend to feel 'squeamish' about the sentence, but may be persuaded to accept the inverse scope more readily than the surface scope. The reason for this ambivalence may be that, firstly, there is the feeling that such sentences are unacceptable due to the PPI 'some' occurring in the c-command domain of its 'anti-licenser' (i.e. negation, which licenses Negative Polarity Items and "anti-licenses" Positive Polarity Items); and secondly, that there is an intuition that the sentence can be 'rescued' and hence is not outright ungrammatical. Thus, its surface scope is acceptable as a denial, if the negation is stressed.

(58) Speaker A: Have you heard? Elmo caught some butterflies.
Speaker B: No, Elmo **didn't** catch some butterflies. He doesn't like to hurt any living things.

Also, the sentence has another acceptable reading, one in which the indefinite outscopes the negation. This reading is harder to access by virtue of its being the non-isomorphic reading. There is substantial evidence for the inverse scope readings being harder; e.g. it takes adults longer to parse sentences under the inverse scope interpretation than under the surface scope for reasons that are not important for the present discussion (cf. Kurtzman and MacDonald 1993 for a review of this literature). What is needed in order to facilitate the inverse scope reading for adults is to create a context that in addition to constructing an appropriate Question under Discussion (by focusing on those butterflies that were not caught), contains a salient group of entities with a unique property, which constitutes a subset that the domain restriction of 'some' may be narrowed to. For example, let's say that the context contains a number of orange butterflies sitting on a hibiscus plant and some yellow and blue butterflies on the rose bush. Let's say that Elmo, whose job is to find all butterflies, succeeds in finding all of the orange butterflies and yellow butterflies, but fails to find any of the blue ones. In this context, the sentence like (ii) repeated below would answer the question in (59) (the italicized part constitutes an implicit

domain restriction):

(59) Did Elmo find all of the butterflies?
(60) No, Elmo didn't find all of them. More specifically, he didn't find some (particular) butterflies, *namely those that are blue and are sitting on the rose bush.*

Previously we observed that for those sentences in which scope reversal would require a quantifier to raise above negation, this option was unavailable. Consistency would require us to rule this option out with respect to 'some' in (ii). However, in this case, just like in (i), a Schwarzschild-style indefinite with a narrowed domain is available. The two readings (the narrow unrestricted and the narrow restricted) paraphrased below:

(61) It's not the case that there exist some x, such that x is a butterfly and x is in C and Elmo found x.
(62) where C={x: x is a contextually relevant butterfly},
(63) C = {x: x is a blue butterfly in the rose bush}.

That the narrow scope reading under negation is an acceptable option for adults was reported in the experiments conducted by Felber 2001. She reports that 5 of her adult subject accepted (and only 2 rejected) the following sentence:

(64) The troll had to eat his hamburger plain because he didn't have some ketchup.

There is additional evidence that the surface scope for sentences of type (ii) is not completely ruled out. Thus, negated sentences with Positive Polarity Items like "some' can be 'rescued' when they become part of a larger negated sentence. Thus (65a) is more acceptable than (65b):

(65) a. It's such a pity that Elmo didn't catch some butterflies.
b. Elmo didn't catch some butterflies.

In (65a) the indefinite is interpreted in the scope of negation without causing as much unacceptability as in (65b) for a reason that is not germane to the question at hand[13]. What is crucial is that the surface scope is not ruled out.

If this sentence is essentially ambiguous, then the explanation for why children fail to access the inverse scope but instead accept the isomorphic interpretation is the same as was

[13] There seems to be variation across speakers in how unacceptable they find this sentence.

proposed for other isomorphic responses for ambiguous sentences, such as (iii) and (iv). Children may have accepted the surface scope because the story-context in which it was presented did not raise an appropriate Question under Discussion. Then, by default, children may assume that the story, asserting that Elmo caught a number of butterflies, addresses the question 'what happened'. Then, they treat the target sentence as a (surface scope) denial of this assertion.

What we can conclude from the discussion of quantifier/negation interactions is that firstly, the paradigm discussed by Musolino is rather consistent. Which sentences lack certain scope readings in adult grammar is not arbitrary, but there are overt structural reasons for why these readings are not available. Therefore, there is no theoretical motivation to posit a developmental stage at which children are limited to the isomorphic interpretation across all scope-ambiguous sentence types in order to avoid the superset problem. Secondly, we can conclude that a failure to access the 'wide scope' interpretation for object indefinites frequently claimed for children, may not stem from the unavailability of such interpretation, but simply because the context does not meet the conditions for such interpretation (e.g. it fails either to raise an appropriate Question under Discussion or to provide a narrowed domain restriction for the specific plural indefinite, or both).

2.3 What should we make of the remaining types of errors?

Now let's consider what conclusion we should make about children's grammatical competence from those studies that demonstrate children's bias towards wide scope or 'referential' indefinites (see table 1). Just as in the discussion on isomorphism, we can also conclude that failing to elicit certain readings under certain experimental conditions doesn't inform us about children's grammatical or lexical limitations. I would like to suggest that these studies inform us about the same type of phenomenon: a strategy children use when they have to interpret indefinites in a vague context, in particular in a context that supplies multiple individuals as a possible domain restriction for the indefinite, and fails to highlight

whether one or all of them are intended to be the member of the restrictor-set. Under these conditions, a child is left with a possibility that one member of the restrictor-set is most relevant. Their error, therefore, can go in either direction – either to construct a singleton indefinite when one was not intended by the experimenter, or to fail to construct one when one was meant. This hypothesis is supported by Su 2000, who compared the child/adult differences in accepting wide scope indefinites in Chinese and English with sentences like 'Mickey Mouse did not ride a dog.' In the story, Mickey Mouse goes to the circus and wants to go for a ride on a horse. After discovering that the horses there were too wild, MM decides to ride a dog. He rides one dog, then another; finally approaches the last dog, but decides not to ride him because he did not look friendly. After the story, the puppet says, "I know what happened. MM didn't ride a dog." As we see, this design suffers from the same limitation with respect to QuD as Musolino's study, therefore it is not surprising that most of the English-speaking children failed to access the inverse-scope reading. What is interesting about this study, however, is that there was a dramatic difference between the performance of the English-speaking children and their Chinese counterparts. While English speaking children differed significantly from English adults in their scope preferences (in the direction of the isomorphic scope), Chinese children mirrored the response pattern of the Chinese-speaking adults more closely (but with a greater bias in the direction of wide scope indefinites).

Table 2 summarizes some of Su's findings:

Table 2. The rate of accepting the indefinite wide scope reading (from Su 2000):

	Adults	Children	Significance
English	50%	33%	$p < .05$
Chinese	64%	77%	not significant

Su explains this discrepancy by different lexical properties of the English *a* and its Chinese counterpart *yi-ge*. While both the Chinese and its English counterpart are

ambiguous between the quantificational and a cardinal ("one") readings, the Chinese lexical item is closer to English numeral 'one'. Chinese children were closer in their response pattern to adults because the lexical item reduced the ambiguity towards the 'singleton' reading and made the task easier. English children accept the wide-scope indefinite reading at a lower rate because *a* can be interpreted as either a singleton or non-singleton. Like the adults, they clearly show the availability of both readings. However, without any special effort of the experimenter to include in the story a relevant question, to which the inverse scope reading would constitute an appropriate answer, and/or to highlight the single relevant individual, children frequently fail to construct the former reading. If we adopt this explanation, we not only can account for a large body of experimental findings, but we also eliminate the contradiction between the so-called 'isomorphism' studies and those studies that show children's non-adult-like responses that are non-isomorphic. A bias towards 'referential' indefinites in the constructions that are ambiguous for adults is children's reaction to the experimental conditions in which the test sentence is presented, and therefore can be pushed in a predictable direction with certain manipulations.

I believe this explanation is essentially correct with respect to the results reported in the third group of studies reviewed in table 1, namely those that show children's failure to access specific indefinites in constructions in which adults strongly favor such readings (e.g. Kramer 1998, 2000). In such studies, children were typically given a set of multiple objects, but were expected to focus on a single one, which they failed to do. Thus, in Kramer 1998, when Dutch-speaking children failed to interpret indefinite subject 'referentially', they were given stories containing multiple protagonists of the sort denoted by the indefinite (e.g. two girls, both of whom were pointed out by the experimenter). The experimenter would show the child a picture of two girls and one boy

and say: "Here you see a girl, a boy, and another girl. They want to do something, but they are not sure what. They may dance, and then again they may do something else." At this point, the experimenter would introduce the second picture, in which one girl and the boy were dancing, while the second girl was doing something else. The experimenter then would point out that one girl was dancing, as was the boy, and one girl was swinging. After that she would "wake up" the puppet, who uttered the target sentence "a girl is not dancing", the indefinite in which is interpreted referentially by Dutch-speaking adults. More than 70% of the 4-5 year-olds rejected at least some of the target sentences and 48% rejected all of them. In contrast, 100% of adults accepted the sentences.

A plausible explanation for this result is that in this experiment children responded in the way adults would, namely to include a single girl in the domain restriction, but not the relevant girl (i.e. not the one that was swinging, but the one that was dancing)[14]. The follow-up experiment reported by Kramer in the same paper supports this explanation. In this second experiment, the second picture (i.e. the one that showed one girl and one boy dancing and the other girl swinging) had a cover with 'windows', which the child had to open to see what each protagonist was doing in order to verify the sentence 'a girl is not dancing'. In this experiment, 100% of the tested adult subjects opened as many windows as was necessary to see if there was a girl that was not dancing, ignoring the girl who was dancing. In contrast, 58% of 4-5-year-olds stopped as soon as they found the first girl whether or not she was dancing and gave their judgment at that point without looking for other girls. So, if they first opened the window behind which the girl was dancing, they would stop looking and give a "no-response".

[14] This was not the conclusion reached by Kramer, who suggested that children failed to get the 'referential' reading and instead interpreted the indefinite 'non-referentially' due to their inability to access a discourse referent.

Presumably, if the first girl they came across was swinging, they would stop looking and give a "yes-response". Interestingly, even the older 5-6-year-old group behaved the same way – 68% of them looked for one girl only. This strongly suggests that children do not lack 'referential' indefinites; but that they fail to choose the same 'referent' for such indefinites as the adults.

The theory that handles the wide scope indefinites by singleton domain restrictions and not through movement would be consistent with both adult and children's data. This approach would also help explain why in Miller and Schmitt's study children were so successful in accessing wide scope indefinites. As I reviewed earlier, M & S purposefully emphasized that the task of the protagonist in the story was to paint all of the eggs, and highlighted the single relevant object among the group – the one egg that was left unpainted. Under these conditions, children had no problem accepting the target sentence like (66):

(66) Kelly did not paint an egg.

3. Conclusion

The mechanism that accounts for wide scope indefinites that I adopted does not require either lexical referential/non-referential ambiguity or syntactic movement. This theory is consistent with the observation that children's biases towards wide-scope indefinites in one group of studies, versus their preferences for the narrow scope indefinites in another group of studies seem to be a function of the pragmatic conditions of the experiment. Those studies that either intentionally or inadvertently set up the situation in such a way that it makes one interpretation salient (e.g. by setting up the story context in such a way that it addresses a question (QuD) that can be best addressed by one of the two available readings or by providing a number of individuals of the sort denoted by the indefinite without focusing on any single one and thus making the 'narrow

scope' reading more salient, or conversely emphasizing one member of the set as most relevant, result in a high rate of responses favoring such interpretation. Those experiments that do not contain clues as to which interpretation is most appropriate in a given situation, expecting the child to be adult-like in making this decision, result in a pattern in which responses are split almost evenly between the possible readings.

One advantage of this approach is that its explanation for children interpretive errors and biases does not hinge on children's knowledge of "QR" or their knowledge of lexical properties of indefinites. Assuming full adult-like grammatical and lexical competence from early on is consistent with the ample studies that show children early acquisition of determiners, such as Schafer and de Villiers 2000; Foley, Lust, Batin, Koehne, and White 2000; Abu-Akel and Bailey 2000, Miller and Schmitt 2003, etc. It is also consistent with the position that children's early competence includes all and only scope configurations permitted by adult grammar. Assuming this theory allows us to account for a large number of empirical observations about the limitations that children exhibit in interpreting scope-ambiguous sentences. Both the observation of isomorphism and observation of the non-referential bias (as well as the referential bias) can be explained in a simple and unified way. Lastly, this theory of children's knowledge of quantification provides an insight into the reason why children frequently make the so-called q-spreading errors under certain experimental conditions. The approach discussed in this chapter puts us at odds with the 'adverbial' theories of quantification in child language on several issues. First, it suggests that the primary cause of spreading errors is not the difficulty children have with the universal quantifier, but with indefinites in environments of scope ambiguity; secondly, it makes different predictions with respect to

availability of various scope readings to children.

In the following chapter, I will consider the varying predictions made by my theory versus those theories that posit a semantic anomaly in children's semantics of quantification. I will then present experimental findings from the study that I conducted in order to test these predictions and show that these results are consistent with the predictions made by the approach that places the non-adult-like property responsible for quantificational errors in children's weak ability to process contextual clues.

In the next chapter, I will present evidence that children may be biased towards the singleton indefinite interpretation even if the context doesn't provide a ready 'reference' for it. I will also show that children make significantly fewer errors with universally quantified sentences in contexts compatible with the singleton (specific) interpretation compared with the non-singleton (distributive) interpretation. In addition, I will show, that the presence on the picture of a perceptually salient single individual in addition to the two overlapping sets causes children to give responses strikingly similar to classic spreading on sentences that do not contain the universal quantifier, but contain another quantifier (namely the negative determiner) in the subject position and an indefinite in the object position.

Chapter 4 Experimental Evidence

4.1 Predictions

I have proposed that instead of focusing on the universal quantifier, as has been done previously, the so-called q-spreading responses should be explained by children's interpretation of indefinites, particularly their choosing singleton domain restrictions influenced by the perceptual salience of the 'extra' object. As I discussed in the previous chapter, my approach relies on the theory of indefinites by Schwarzschild 2001, in which indefinites are always quantificational, but can appear referential simply because of their ability to have singleton domain restrictions. Singleton domain restrictions affect the semantic scope relations independently from any syntactic mechanism of reversing scope relations (i.e. a covert movement operation or "QR", which alters the surface c-command relations in order to determine the relative scope of the scope bearing elements in the structure). In particular, those indefinites that combine with singleton domains, even when syntactically taking narrow scope and therefore normally receiving the non-specific interpretation, would have the appearance of being 'wide scope' or 'referential' and essentially render the sentence scope-neutral (in Schwarzschild's terminology), thus obscuring the surface syntactic scope relations.

An additional issue raised in the literature on the acquisition of quantification is whether syntactically inverse scope readings are available to children from early on, which has been treated in most studies as an issue independent from children's knowledge of semantics of quantification. Because the mechanism that accounts for specific indefinites that I adopted does not require syntactic movement, the issue of whether or not children have adult-like knowledge of "QR" remains a tangential

question, a positive answer to which is not required in order to claim children's full semantic knowledge of universal quantification. However, this issue is relevant for the 'spreading' debate because, while the Full Competence hypothesis is consistent with the claim that children know all and only those scope readings that are allowed in adult grammar, "adverbial" theories of quantification theories, such as Event Quantification (Philip 1995), maintain that at the initial stage, universally-quantified sentences should obligatorily have universal wide scope since in that theory the universal quantifier occupies the position in which it scopes over the entire sentence. It has even been suggested that the basic underlying reason for q-spreading is children's difficulty with QR. Thus, Philip (1995) suggests that the reason why children interpret the universal quantifier as an adverb instead of a determiner may be that the children have difficulty with the operation of QR. According to Philip, applying QR may be problematic for children since it is an abstract relation (since the movement occurs post-syntactically) and because it results in establishing a discontinued dependency between the moved constituent and its trace. This view is close to the maturation hypothesis of Borer and Wexler (1987), who claimed that children developmentally lack the mechanism of processing A-chains and consequently have difficulty comprehending derived passives. In Philip's view, because of their difficulty with adult LF, children resort to event quantification, which is not derived via QR, but with a mechanism described in Heim's terms as a "Quantifier Construal Rule" (Heim 1982), and does not require maintenance of an abstract discontinuous dependency. As Philip put it, "all that matters for semantic interpretation of the universal quantifier in such a representation is that this quantifier has scope over the entire sentence; its original position in the surface structure is completely

irrelevant" (p.50).

Hence, under this theory, sentences (1) and (2) for children should have identical semantic form derived from an LF in which the universal takes scope over the whole sentence and hence over the indefinite.

(1) Every farmer rode a donkey.
(2) A farmer rode every donkey.

This predicts that any unpaired individual – a non-donkey-riding farmer or a donkey not ridden by a farmer should equally result in a spreading response for either (1) or (2). Furthermore, if the two sentences have the same LF, distinct from its surface syntax for both of them, there is no obvious reason why the rate of q-spreading should be different for the two sentence types (apart from the possibility that the greater distance which 'every' has to move in (2) may affect the rate of errors). In addition, if the adverbial theory is correct and children always require distributivity and complete co-variation between farmers and donkeys for both (1) and (2), both should be rejected in collective contexts, where all farmers ride the same donkey or all donkeys are ridden by the same farmer, as frequently or more frequently as in distributive contexts, as long as there are unpaired farmers or donkeys.[1]

In contrast, my approach makes different predictions. If children have adult-like knowledge of quantification, we must at least allow a possibility that they have both surface and inverse scope interpretation for both sentences, given that pragmatic conditions in which the test is administered do not inhibit some reading. In experiment 1, I will test this prediction with respect to the inverse scope for sentences like (2).

---

[1] For (2), children presumably should allow for one specific farmer to ride all of the donkeys, but only in case there are no extra farmers. Similarly, for (1) the collective context with no extra donkeys should result in a "true" response, but a donkey would falsify it.

Normally, when sentences of this type are used in "spreading' studies, children are shown a picture with a number of farmers each riding a different donkey and an extra farmer. When, in this context a child rejects (2) as false and points to the extra farmer, it is taken as evidence that she assigned the sentence universal wide scope reading, in accordance with the event quantification semantics, and the unpaired farmer under such reading made the sentence false for the child (Philip and Aurelio 1991, Philip and Takahashi 1991). However, since in these studies there are no controls insuring that the sentence does in fact receive universal wide scope reading for children, we cannot conclusively say what causes them to reject the sentence in that context. It is quite plausible that they reject the sentence because they interpret it under the surface scope and look for one specific farmer to be riding every donkey, and since the context does not contain such a farmer, they reject the sentence. This is especially plausible given the preponderance of experimental evidence showing that the inverse scope reading is harder to elicit from children than the surface scope and that the former can be facilitated by more carefully designed experimental conditions (Hulsey et al. 2005). Therefore, it is important to tease apart the surface and the inverse scope interpretation of sentences like (2) and to study whether there is a difference in "q-spreading" rate for each reading in a better controlled environment. I will address this issue in experiment 1.

If children in fact prefer the surface scope for (2), we can expect to see a difference in the rate of 'q-spreading' responses between (2) presented in the context consistent with its surface scope interpretation (e.g. one farmer-multiple donkeys group plus extra farmers) and (2) presented in the context consistent with its inverse scope reading (a number of farmer-donkey pairs plus an extra donkey). Children should reject

(2) significantly more frequently if it's presented in a distributive context than when (2) is presented in the context consistent with the group reading. Furthermore, just as the presence of extra farmers shouldn't matter for increasing errors for the group contexts, the absence of an extra donkey shouldn't matter in reducing the rate of errors for distributive contexts. In other words, I predict that sentence like (2) would be accepted by children at a high rate in "group" contexts regardless of the presence of extra agent-type individuals, and rejected by children at a relatively high rate in a distributive context even if the context contains no extra patient-type individuals. This prediction will be tested in experiment 1, discussed in the next section, where I will show that this prediction is borne out.

There is another scope related issue that needs to be investigated. If children have adult-like knowledge of quantifier scope, as I propose, one element of this competence should be their knowledge of asymmetric entailment relations between the surface and inverse scope readings in sentences like (1). In this example, the wide-scope indefinite reading entails the wide scope universal reading. This means that if the sentence is true under the wide scope indefinite reading (which says that there is a donkey such that every farmer rode it), it would necessarily be true under the wide scope universal in the same context as well (which says that for every farmer x there is a donkey y such that x rode y).

However, in those contexts that make the sentence true under the universal wide scope reading, the sentence would be false under the indefinite wide scope construal. This amounts to a situation, when (1), presented in a context consistent with the wide scope indefinite reading, is unambiguous – true under either surface or inverse scope

configuration. When, however, it is given in a distributive context, the sentence is ambiguous – either true if construed under the surface scope, or false if construed under the inverse scope. The same applies to (2), for which the sentence is true under both readings if presented in the 'wide scope indefinite' situation (henceforth 'collective' contexts), but if presented in the 'narrow-scope-indefinite' situation (henceforce 'distributive' contexts), the sentence is true only under the 'distributive' reading and false under the 'collective' reading.

If my hypothesis is correct in claiming that q-spreading is children's reaction to the sentence ambiguity and the use of the perceptual salience of the extra object in resolving the ambiguity, this state of affairs should result in a distinct error rate for 'collective' vs. 'distributive' asymmetric contexts. Since in a collective situation, where every farmer rides the same donkey, there is no possible interpretation of (1) that would give a "false" response in that situation, children's decision of which reading to select should not be affected by the visual asymmetry and hence, they should exhibit a low rate of q-spreading errors. This experimental situation should in effect be analogous to those tests that disambiguate the sentence by presenting it in a rich story-context. On the other hand, children should exhibit a higher error rate with sentences like (1) in distributive asymmetric contexts, particularly in the single extra object condition since in this situation they are giving truth-value judgments in a situation that requires them to resolve ambiguity. This prediction will be tested in experiment 2.

Finally, if the main culprit in q-spreading is not the universal quantifier, but the indefinite, we can ask ourselves whether children will respond in a "spreading"-like manner to sentences containing an indefinite and a quantifier of another type, instead of

the universal, given that the sentences are presented under the typical "spreading" inducing conditions, e.g. visual context with a salient single "extra" individual. The approach advanced here predicts that spreading-like errors can indeed be elicited using sentences without the universal quantifier if such sentence contains an indefinite and the context contains a salient single individual to be a potential candidate for a singleton domain restriction for the indefinite. We also predict that the error rate can be manipulated by controlling the degree of visual salience of the extra individual. This prediction will be addressed in experiment 3.

4.2 Experiment 1

4.2.1 Goal, Participants, and Method

Goal: To test whether English-speaking children can correctly interpret sentences like (3) under both surface and inverse scope and whether their success in this task is affected by conditions of contextual asymmetry between the number of individuals in the agent and theme sets and and/or satisfaction of the Condition of Plausible dissent.

(3) A farmer milked every cow.

Participants and Materials

The participants were 21 English-speaking preschool children (median age 4; 6) recruited at the University of Connecticut Child Labs. Five children, who failed the training, were excluded, bringing the sample to 16 children. We also tested 16 adult controls, who were University of Connecticut undergraduate students. The adults were given the same materials as the children, but in a paper and pencil form – the written stories were presented in the same randomized order as the children, and each was followed by the target sentence, which the subjects were asked to judge as true or false.

Participants[2]

Child	age
Mike	3;4
Nell	3;6
Valerie	3;10
Billy	3;10
Lionel	3;11
Cathy	4;2
Amy	4;3
Tatiana	4;6
Mathew	4;6
Steve	4;7
Ron	4;7
Ivan	4;7
Jake	4;9
Greg	4;9
Jane	4;10
Hellen	4;11
Eva	4;11

The method used with children was the standard Truth-Value Judgment Task (Crain and Thornton, 1998). The materials consisted of sentences like (4) presented in contexts that satisfied one of the following three conditions:

(i) compatible with the inverse scope, but not with the surface scope;
(ii) compatible with the surface-scope, as well as the inverse scope by entailment;
(iii) incompatible with either the surface or inverse scope

(4) A girl tasted every cake.

The two possible interpretations for this sentence are given in (15) and (16):

(5) $\forall y$ [cake(y) $\rightarrow$ $\exists$ x [girl(x) and x tasted y]]
'For each cake y, there was a (possibly different) x, such that x was a girl and x tasted y.'

(6) $\exists x$ [girl(x) & $\forall y$ [cake(y) $\rightarrow$ x tasted]]
'There was an x, x is a girl and x tasted every cake.'

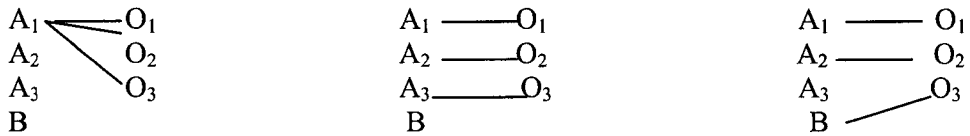
Under the scenario of type (i), compatible only with the reading in (15), the story contained three girls, who each tasted a different cake. In a situation of type (ii),

[2] Children's names in the table are pseudonyms.

compatible with both readings, there was one girl who tasted every cake, while other girls did not. Finally, in a situation of type (iii), there was neither a single girl who tasted every cake, nor an exhaustive pairing between girls and cakes.

All scenarios satisfied the condition of Plausible Dissent by containing an alternative individual (e.g. in the above-mentioned scenario, three girls and one boy). The alternative character always considered performing the activity described in the story, but failed to do it.

7a. Surface scope context: b. Inverse scope context: c. False on both readings:



Each child was given four items of each type and four fillers. The testing was divided into two sessions. Each test item was pronounced with a focus-neutral intonation, without stressing either the subject or the object NP. The stories were acted out with small toys by the experimenter, before the test sentence was presented.

Sample surface scope, inverse scope, and false stories with their respective test sentences are given below:

(8) **Sample surface scope story**:
Once there were three dragons: small, big, and extra-large. They were friends with a lion. One day they decided to have some fun. What can be more fun for dragons and lions than chasing a knight through the woods! So they went in search of a knight. First, it was the small dragon's turn. He saw a green knight and roared, but the green knight was not afraid. He said: "Oh, what a cute little dragon! Come I'll scratch behind your ears!" The little dragon hated to be treated like a pet! He ran away. Then the big dragon took a turn. He found a black knight with a sword. When the knight saw the dragon, he said: "Great! A decent size dragon! Hey dragon, come out and fight!" But the dragon didn't want to fight a guy with a sharp sword. Next, it was the lion's turn. When he saw a red knight, he roared at him. But the red knight said: "Wow! A lion! I will take him to the zoo." Of course, the lion had to run away because he hated the zoo. Then the extra-large dragon

said: "Don't worry. I will show them." He knew how to breathe fire. He scared the red knight, and the black knight, and the green knight.

Puppet: A dragon scared every knight.

(9) Sample inverse scope story:

There were three sisters-princesses and their brother, the prince. They liked to read fairy-tales. Once they read a story in which a princess kissed a frog and the frog turned into a handsome prince. When they went for a walk in the palace garden, they saw three handsome looking frogs. The youngest sister said: "I am going to kiss the yellow frog. I am sure it will turn into a prince." And she did. But the frog remained a frog. The middle sister said: "I think the red frog really is a prince. I will kiss him and see." And she did. The frog did not turn into a prince. The oldest sister said, " I am sure the green frog is really a prince. I don't feel like kissing it, though. It looks so slimy. Can you, brother, kiss it for me, please?" But the brother said "Oh, no, only a princess, not a prince, can break the spell.' So the oldest sister had to kiss the green frog herself. Alas, it didn't turn into a prince either.

Puppet: A princess kissed every frog.

(10) False:

Once there was a wicked pirate. He caught three dogs, put them all in cages and hid them in his house. The dogs were very sad because they missed their families. A rescue team came to help the dogs: three bears (a brown bear, a black bear, and a purple bear), and a space alien. The brown bear opened the first cage and set the first dog free. The black bear opened the second cage and set the second dog free. The purple bear said: "I will open the last cage." But the last cage was very tricky, and the little bear couldn't figure out how to open it. The alien said: "Let me try." He was very smart and he opened the last cage and set the last dog free.

Puppet: A bear opened every cage.

Each condition contained four items, pronounced with a focus-neutral intonation (12 test items in total), plus fillers. Prior to testing, each child received 4-6 training items. Each training item consisted of an unambiguous sentence containing one quantificational element only (the universal *every* or indefinite *a(n)* in either the subject or object position, as is illustrated in (11) - (13). The initial training contained two false and two true items.

(11) Every child drank milk.
(12) A boy climbed the tree.
(13) The brown pony carried every child across the meadow.

The criterion for inclusion in the study was correct acceptance and rejection of 4 training sentences.[3] A sample of a training story for which a false answer was expected is given in (14):

(14) Three children, two girls and one boy, were playing together on the playground. They got very thirsty and decided to go home to get something to drink. In the refrigerator they found some milk and some soda. The youngest sister said, "Mm... Nice cold milk! That's what I will drink.' And she drank a glass of milk. Her brother said, " I am going to drink some soda." And he did. The oldest sister said, " Soda is not good for you. Milk is much better." And she drank some milk.

Puppet: Every child drank milk.

The stimuli were presented in a fixed random order.

4.2.2 Findings

Table 2 below shows how many sentences were accepted out of the total 4 for each type. Table 3 summarizes the results.

Table 2 **Number of Accepted Sentences in Each Condition**

	surface	inverse	false	fillers
Mike	3	0	0	3
Ivan	4	4	0	4
Jake	3	4	0	4
Tatiana	3	0	0	4
Amy	4	4	1	1
Ron	4	0	0	4
Cathy	2	0	0	3
Billy	4	3	1	3
Valerie	3	2	0	3
Lionel	1	4	1	4
Jane	3	0	0	3
Mathew	4	3	1	4
Hellen	3	4	0	4
Eva	3	0	0	4
Nell	4	0	0	3
Steve	2	4	0	4

[3] All children were given 4 initial training items. Those who passed all of them were not given any additional items. Children who failed 1-2 of the initial items were given additional 2 items of the type they missed. Of those, children who failed at least three in total were excluded.

Distribution of Responses in Children (N = 16)

	accepted ≥3	rej ≥3	split
Surface	13	1	2
Inverse	8	7	1
False	0	16	0
Fillers [4]	15	1[5]	0

Table 3

As we can see, the children are nearly unanimous in their acceptance/rejection of certain sentence types, but they are split with others: while children were adult-like in almost uniformly rejecting false sentences and accepting the surface scope, they fall into two almost equal groups with respect to inverse scope.

The main goal of the study was to obtain more fine-tuned findings about the rate of "q-spreading" with respect to sentences like (4) by teasing apart the two scope readings available for such sentences. Our results show that indeed there is a sharp difference in the rate of incorrect rejections with respect to the two readings. As we predicted, the test sentences were much more likely to be rejected in 'distributive' contexts, despite the complete symmetry between the agent and theme sets of protagonists.

The surface scope contexts contained an asymmetry. Thus, in the story shown in (8), out of the three dragons present in the scenario, only one scared the three nights, while two others did not. Recall that such asymmetry under the 'adverbial' theory is expected to cause children to reject the sentence and the children are expected to explain the

---

[4] Since fillers varied between true and false, the table indicates correct responses in the acceptance column and incorrect in the rejection column.

[5] The child who did poorly on the fillers (Amy) was tested first. After doing the experiment with her, we adjusted the fillers to eliminate unnecessary complications in the story, which had to do with the plot and not the structure. Since Amy was very consistent in rejecting the false sentences while accepting both surface and inverse scope ones, we decided not to exclude her from the result table. However, eliminating her from the results would not change the general pattern.

rejection by pointing to the extra dragon(s). Our results indicate, however, that when the context was presented in a rich story form, instead of a schematic picture showing dragon/night pairs, the asymmetry did not cause children to reject the sentence. The children consistently accepted stimuli in 'surface scope' contexts, demonstrating that they allow wide-scope indefinites, and when the sentence is interpreted that way, asymmetry in the context does not result in errors.

In the 'inverse scope condition', on the other hand, the contexts were symmetrical – e.g. in the sample story given in (9), there is an exhaustive pairing between the princesses and frogs that were kissed by the princesses. Despite the exhaustive pairing between the agent and theme sets, this condition presented difficulty for some children. The children in our sample were almost evenly divided: 8 out of 16 children accepted the stimuli in the inverse scope contexts and 7 rejected them (1 remaining child gave an equal number of yes- and no-answers).

Those who accepted the inverse scope support our hypothesis that adult-like knowledge of scopal ambiguities is available to children from early on. This experiment does not allow us to rule out the possibility that the difference between the "inverse" and "no-inverse scope" groups amounts to a developmental difference between those who have matured into a stage at which children allow inverse scope vs. those who remain at the isomorphic stage. However, as I have discussed in chapter 3, the Observation of Isomorphism (Musolino 1998) merely reveals a strategy that leads the child to choose the isomorphic interpretation given certain pragmatic conditions of the test, and does not describe children's competence as lacking the inverse scope (Gualmini 2003, Hulsey et al. 2004). In light of the latest findings with respect to the "Isomorphism" hypothesis, I

will adopt the assumption that inverse scope is generally available to children, but is difficult to access and requires certain conditions to be met by the task in order to make the inverse scope reading more salient. We did not address this issue in my experiment, which is presumably why a number of children rejected the inverse scope reading. However, I believe that those who rejected the inverse scope still present an interesting finding.

Firstly, as I noted earlier, these children rejected the sentence in the inverse-scope context despite the fact that this context exhibits symmetry, supporting my claim that children's quantificational errors are not fully contingent on the contextual asymmetry. Secondly, the explanations offered by the children for why they rejected the sentence in the inverse scope contexts, were quite revealing. All of such children explained that the puppet was wrong when he said that *a princess* kissed every frog because in the story there were *three princesses, not one.*[6] This explanation for rejecting the sentence was consistent with my hypothesis that typical 'quantifier-spreading' errors involve assigning a singleton restriction to the indefinite.

Adult controls were unanimous in their acceptance of the surface scope, but, interestingly, were also somewhat divided in their acceptance of the inverse scope: 3 out of 16 adults rejected three or more of the items in the inverse scope condition (see table 4 for the summary of the results).

---

[6] Most children's explanations consisted of repeating the story line with the emphasis on the multiple members of the subject set (e.g. **this** princess kissed **this** frog, **this** princess kissed **this** frog, ***and*** **this** princess kissed **this** frog). Some, however, were less oblique. Thus, one child explained why she said 'No' by saying, 'He [the puppet] thought there was one [princess]'.

Adult Controls (N = 16)

	accepted ≥3	rej ≥3	split
Surface	16	0	0
Inverse	13	3	0
False	0	16	0
Fillers	16	0	0

Table 4

Since even some of the adult native English speakers rejected the target sentences in the contexts in which others accepted them, it seems reasonable to conclude that what we found was a failure of these subjects to access this interpretation rather than a lack of this reading in their grammar. Such a view is supported by the studies showing a general preference for surface scope over the inverse scope in monolingual adult English-speakers (Kurzman and MacDonald 1993, Pica & Snyder 1994).

In K&M's study, in a self-paced reading task, subjects had to judge whether a sentence with two quantifiers (*a* > *every*) was compatible with a disambiguating continuation. While over 80% of continuations compatible with the surface scope were accepted, only around 30% of those compatible with the inverse scope were accepted. K&M term this preference 'the single reference principle'. They explain this preference by saying that single reference is simpler to represent than multiple reference, it is always possible and often obligatory (as when *a* is an only quantifier), and to switch from single to potential multiple reference involves a complex alteration of semantic representation (Kurtzman and MacDonald, 1993). Thus, they view the 'single reference' bias in adults as a sort of a garden path and not a grammatical phenomenon.

A final observation from our results was that satisfying Plausible Dissent was not a sufficient condition for some children's acceptance of the target sentences in the inverse

scope context. The condition of Plausible Dissent was satisfied across all of the experimental conditions, which, however, did not result in a uniform acceptance of the inverse scope sentences.

4.3 Experiment 2

4.3.1. Goal, Participants and Procedure

Goal: to test the prediction that children know the asymmetric entailment relations between the surface and inverse scope and consequently would make fewer "q-spreading" errors in "collective asymmetric" contexts compared with the "distributive single asymmetric" contexts.

Participants: 15 preschool and kindergarten children attending University of Connecticut Child Labs, median age 5; 4.

Materials and Procedure: universally quantified sentences presented in one of the following contexts:

i. distributive with a single extra item
ii. collective with a single extra item

The contexts were presented as pictures, samples of which are given below:

a) distributive with one extra item

Every baby is inside a flowerpot. ($\forall < \exists$: true; $\exists < \forall$: true)

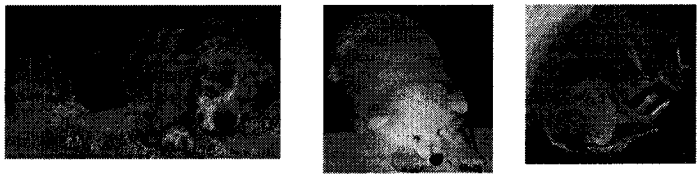
c) collective with one extra item

Every girl is pushing a car. (∀ < ∃: true; ∃ < ∀: false)

Each condition consisted of 4 items plus fillers, presented in a fixed random order. The child was shown a picture, after which the puppet uttered a test sentences. The task of the child was to judge whether the puppet's statement was true of the picture. In addition to the three test conditions, children were given the following control conditions:

a) "no-controls": "one cat is eating, but both cats are sleeping":

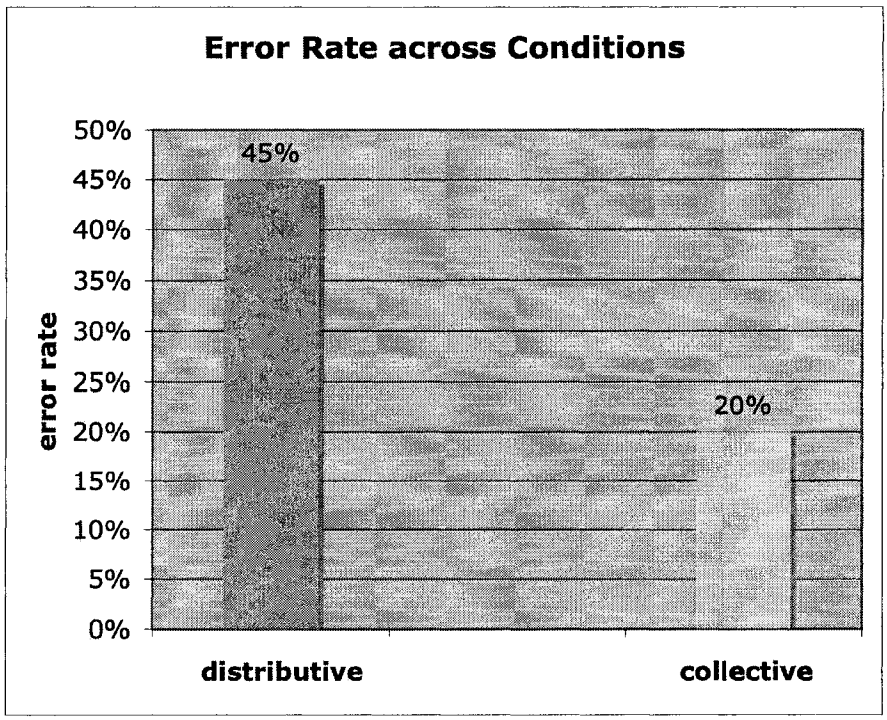
b) "yes-controls" : "Every puppy is sleeping":



4.3.2. Results

As I predicted, we found a distinct rate of error across the three experimental conditions with the distributive with a single extra item resulting in the highest rate of error, and the collective with a single extra item in the lowest rate of errors. The findings are summarized in the following chart:

figure 1:



Paired t-test: t(13)=3.61, two-tailed p=.003

We can see that there is a highly significant contrast in the rate of errors in the distributive contexts with a single extra item and the collective contexts with a single extra item: the former resulted in a rate of errors more than twice that in the latter condition despite the presence of a single extra item in both types of context. This observation is in concert with the hypothesis attributing 'classic spreading' to a pragmatic strategy used by children to resolve an ambiguity in contexts that have a single extra individual because its presence suggests to the child that a singleton reading for the indefinite should be considered and therefore resulting in a certain percent of rejections. According to this explanation, the extra individual triggers errors in conjunction with a number of agent/theme pairs because such context is maximally ambiguous: it contains

the two sets of individuals across which the event is distributed suggesting the distributive reading, and the single individual of the sort denoted by the indefinite, suggesting the singleton indefinite reading. In this situation the child has to make a decision about which reading the speaker intends for the sentence. Consequently, the responses are not limited to either 'true' or 'false', but are split between both. On the other hand, in the 'collective' context, the ambiguity is minimized because, as we discussed earlier in this chapter, the indefinite wide scope reading entails the universal wide scope reading; the two readings are essentially indistinguishable since the only possible judgment for both of them is "true". Consequently, presence of an extra item does not matter to the same extent as in the 'distributive' contexts because it doesn't stand for a potentially different reading. Another final observation (not included in the graph) is that children's responses were nearly 100% on both the "no- and the "yes-fillers", demonstrating that the responses were not random.

4.3.3. Additional support

One major goal of the two experiments presented above was to refute the claim that for children sentences like (1) and (2) have identical LF. In this section, I would like to review some independent evidence that children are sensitive to the syntactic position of the universal quantifier. Meroni et al (2002) showed that children possess knowledge of the asymmetry between internal and external arguments of the universal quantifier with respect to the property of Downward Entailment (DE) (allowing inference of the truth of a proposition predicated of a set to its subset). In adult language, the two arguments of the universal quantifier *every* differ with respect to this property. The

quantifier is downward entailing only in its internal but not in its external argument, i.e. an inference from a set to its subset is licensed only for the internal argument of the universal quantifier, as in (15), but not the external one as in (16):

(15) Every [boy who rode an elephant] [fell off] → Every boy who rode a wild elephant fell off.

(16) Every [boy who fell off] [rode an elephant] *→ Every boy who fell off rode a wild elephant.

Another diagnostic for DE is the licensing of NPIs, such as *any*, which are allowed only in the internal argument of *every*, as in (17), but not the external one, as in (18):

(17) Every [boy who rode any elephant] [fell off].

(18) *Every [boy who fell off] [rode any elephant].

Another consequence of DE is that the disjunction operator *or* can receive a conjunctive interpretation in the internal, but only the exclusive disjunctive interpretation in the external argument of *every*.

(19) Every [boy who rode an elephant or a zebra] [fell off] ⇔ Every boy who rode an elephant fell off and every boy who rode a zebra fell off.

(20) Every [boy who fell off] [rode an elephant or a zebra] *⇔ every boy who fell off rode an elephant and every boy who fell off rode a zebra.

Meroni et all, 2002 tested whether English-speaking children know the difference between the internal and external arguments of *every* with respect to the licensing of

conjunctive *or*.[7] They report that children know that *or* can be used conjunctively in the internal argument of *every*, but only as exclusive *or* in the external argument.

As a further argument for the view that claims that children are sensitive to the syntactic position of the quantifier, it was reported that the rate of children's errors is affected by whether the universal is in the subject (henceforth right -spreading error) or object (henceforth left-spreading error[8]) position. Thus, H-K Kang (in press) reports a significant difference between the rate of the right-spreading (51%) and the left-spreading error (73%) in English-speaking children[9].

One other interesting piece of evidence that children are sensitive to the syntactic position of the universal quantifier comes from the study of quantification in child Catalan (Gavarro and Escobar 2002) and Spanish (Escobar 2002). In their study, G&E were testing the Event Quantification hypothesis with respect to sentences like (21) and (22). In both Spanish and Catalan, a grammatical property (precise nature of which is not crucial for the question at hand) would require sentences like (21) to have frozen scope, where the universal quantifier takes wide scope. When children were asked to judge the truth value of such unambiguous sentences in the contexts like those outlined in a and b, their responses closely mirrored those of adults, and they accepted sentences like (21) in

[7] The knowledge of this property was also tested by Gualmini et al, 2001, who tested whether children allow conjunctive use of *or* in the internal argument of *every*. They report children's competence in this respect.

[8] The term left-spereading and right-spreading error indicates the syntactic position of the universal, which the child is claimed to fail to restrict by the common noun, but instead 'spreading' it to the whole sentence. If it is in the subject position the error is called right-spreading, and if it is in the object position it is called left-spreading.

[9] In this experiment, children were shown a picture, asked to describe what they saw, and then presented with a Yes/No test question. The description by the child of the picture consisted of naming the 'characters' and the activity they were engaged in, e.g. This bear is holding a honey pot, this bear is holding a honey pot, this bear is holding a honey pot too, this honey pot is alone... This served the purpose of testing weather the child understood the picture correctly. However, it did not provide an enriched context and contrary to the author's claim did not satisfy the Condition of Plausible dissent.

the contexts analogous to (a) despite the 'extra' individuals, which typically elicit 'spreading' responses. At the same time, they correctly rejected the sentence in the context of (b).

(21) Is an elephant carrying all the balloons?

Contexts: a picture showing

a)	b)
$elephant_1$ carries $ballon_1$, $balloon_2$ and $balloon_3$	$elephant_1$ carries $ballon_1$
$elephant_2$	$elephant_2$ carries $ballon_2$
$elephant_3$	$elephant_3$ carries $ballon_3$
boy	boy

In G&E, 94% of the children accepted the surface scope reading (situation a) and 94% of 3- and 4-year olds (as well as 100% of older children) rejected the inverse scope reading (situation b). If in these children's grammar, the universal quantifier was an adverbial quantifier over events, we would expect to find an opposite: children should have rejected the sentence in (a) since it has elephants not carrying balloons, but accepted it in situation (b), since it exhaustively pairs elephants and balloons.

Interestingly, we find a different pattern of responses for G&E's children when the test sentences have a universally quantified subject and indefinite objects. When children were asked to judge the truth-value of (22), their responses across all of the experimental conditions had a higher error rate than what we saw with respect to (21).

(22) Are all the elephants carrying a balloon?
Context: picture showing

a.	b.	c.
$elephant_1$ carries $ballon_1$	$elephant_1$ carries $ballon_1$	$elephant_1$ carries $balloon_1$
$elephant_2$ carries $ballon_2$	$elephant_2$ carries $ballon_2$	$elephant_2$ carries $balloon_2$
$elephant_2$ carries $ballon_2$	$elephant_3$ carries $ballon_3$	$elephant_3$ carries $balloon_3$
boy_1 carries $ballon_4$	$elephant_4$ carries $umbrella_1$	$horse_1$ extra $balloon_4$

The younger group (3-4-year-olds) gave responses that were only 78% correct on situation (a), which typically gives rise to 'perfectionist ' errors in English-speaking

children, 56% correct on (b), typically causing the so-called 'under-exhaustive' error, and 72% correct on (c), a "classic spreading" context. Results in the Spanish study by Escobar were very similar to the Catalan results by G&E.

I believe that these results can be easily accommodated under my approach, since in (22), but not in (21) there is an ambiguity in how the indefinite can be interpreted. In (21), the indefinite is obligatorily specific and children make very few errors judging the truth-value of this sentence regardless of the presence of extra individuals in the picture. In (22), on the other hand, one may interpret the indefinite object as either specific or non-specific depending on how its domain restriction is construed. The ambiguous stimuli trigger a higher rate of errors than the non-ambiguous ones. In contexts of type c, the presence of an extra single individual triggers a certain percentage of 'singleton' responses. I will discuss the 'perfectionist' and the 'under-exhaustive' responses later on.

Thus, we can account for the puzzling contrast in the rate of errors Spanish and Catalan-speaking reviewed above if we take into consideration a lack of scopal ambiguity in those sentences that produced very low error rate and a presence of scopal ambiguity in those that produced a high error rate. Furthermore, this approach explains the errors not by appealing to children's lack of certain semantic knowledge, but by their full semantic competence, which leads them to recognize the potential ambiguity and, presumably for reasons independent from their linguistic competence, handle the ambiguity resolution differently from adults.

There is independent evidence that Spanish children have full knowledge of the semantic properties of indefinites, which comes from the study conducted by Miller and Schmitt (2003). The goal of their study was to see whether Spanish-speaking children

are able to have both specific and non-specific interpretation of indefinites under negation. They studied Chilean Spanish-speaking children's responses to sentences containing indefinites, which are ambiguous, as in (23), and bare singulars, which are obligatorily non-specific, as in (24):

(23) El niño no se compró un perro. (neg > a; a > neg)
The boy neg rfl bought a dog
"The boy didn't buy a dog."

(24) El niño no se compró perro. (neg > a; *a > neg)
The boy neg rfl bought dog
"The boy didn't buy a dog."

When children were asked to judge the truth-value of sentences like (23) and (24) in the situation compatible only with the inverse scope (e.g. the protagonist bought a number of dogs but one), there was a significant difference in the rate of 'false' judgments between (23) and (24). While with respect to the unambiguous (24), the rate of 'false-judgments' was as high as 76%, it was a low 36% with respect to the ambiguous (23). In comparison, adults' 'false-responses' were 99% and 43% for the bare singular vs. indefinite respectively, the pattern the children essentially mirrored. It is plausible then that G&E's children's 'spreading' responses were due to specific indefinite readings intruding into how they interpret (22) and causing them to make the errors.

In short, the data reviewed in this section, as the results of the experiments presented in the previous two sections shed some light on the mechanism that creates 'spreading' errors: On the one hand, they show that children have sophisticated semantic knowledge of quantification (including the knowledge of very subtle properties as shown by their differential treatment of the internal and external arguments in a universally quantified sentence with respect to the property of Downward Entailing and their

sensitivity to a potential ambiguity due to scope interactions between the universal quantifier and indefinites). On the other hand, we can see that children's weakness (presumably lying outside of their linguistic competence and underlying their weak ability to compensate for insufficient context), makes them prone to errors when context is required in dealing with the complexity created by their sophisticated grammar.

4.4 Experiment 3

4.4.1 Goal, Participants, and Method

If our hypothesis is correct in maintaining that the main culprit in q-spreading is not the universal quantifier, but the indefinite, we can ask ourselves whether children will respond in a "q-spreading"- like manner to sentences containing an indefinite and a quantifier of another type, instead of the universal, given that the sentences are presented under the typical "spreading" inducing conditions, e.g. visual context with a salient single "extra" individual. Another, related, aim of this experiment is to gain a better understanding of the role visual asymmetry plays in inducing the so-called q-spreading errors. The experiments that report high rates of q-spreading typically present target sentences paired with a picture showing participants and their schematic relationships with each other in the event referred to in the target sentence (e.g. boys riding ponies). Typically there is a distribution of participants across sub-events, but a lack of one-to-one correspondence between the members of the restrictor set of the universal quantifier and the members of the set denoted by the common noun of the second argument. Why such contexts result in a high rate of q-spreading errors has been widely debated in the literature. As discussed in chapter 2, previous approaches tied children's interpretation of the quantified sentence and contextual asymmetry by positing semantics, in which the

universal quantifier fails to be correctly restricted and 'spreads' to other arguments at some grammatical level. This mechanism of 'quantifier spreading' has been placed at the level of syntax (Roeper et al. 2004), semantics (Philip 1995), or syntax-to-semantics mapping (Geurts 2004). Since my proposal posits a different role for the visual asymmetry in triggering children's non-adult-like responses, namely tying it to how the domain restriction for the indefinite may be constructed, we predict that errors that look strikingly similar to classic quantifier-spreading errors can be elicited with sentences that instead of the universal quantifier contain the negative determiner.[10]

Goal: to test the following research hypothesis:

Hypothesis: "Classic spreading" responses result from children's strategy of using asymmetry in the visual context as a pragmatic clue for selecting domain restrictions for the quantified phrases. If the picture contains a salient single individual with the property denoted by the indefinite (i.e. a salient single wagon in case of a sentence "no boy is pulling a wagon"), children are likely to see its salience as an indication of its relevance for the speaker and construct the domain restriction for the indefinite to include exactly one individual (the extra wagon). Consequently, the sentence 'no boy is pulling a wagon' would receive "wide scope" indefinite interpretation, true in the situation depicted in the scenario in which some boys are pulling wagons, but one salient wagon is not being pulled.

Research Questions:

- Can q-spreading-like errors be elicited with sentences containing the negative determiner, given that the sentence contains an indefinite, and the context contains a salient single individual to be a potential candidate for a singleton domain restriction for the indefinite?
- Can the error rate be manipulated by controlling the degree of visual salience of the extra individual?

Participants

Our sample consisted of 19 preschool age children (3- and 4-year old) attending UConn Child Labs. After the training session, 3 of them were excluded after they failed the

[10] I am indebted to Y.Sharvit for suggesting the sentences with the negative determiner as a good test case for my hypothesis.

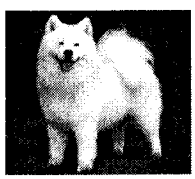
training, bringing the sample to 16. Five adult control subjects have also been interviewed.

Method

I used a version of the standard Truth Value Judgment Task. Children were shown pictures and were asked to judge the truth-value of target sentences like those in (25)-(27). The variable property across two of the experimental conditions I will present first was the degree of visual salience of the extra individual in the object set - single extra object and non-salient multiple extra objects. Each experimental condition consisted of 4 sentence/picture pairs presented in a fixed random order plus fillers. No linguistic context was given (no background story) to replicate the contextual paucity that typically gives rise to q-quantifier-spreading errors. Prior to presenting the target sentence, children were asked to point out who was in the picture to ensure its full comprehension. After that, the puppet uttered the test sentence, and the child was asked to judge whether the puppet's description of the picture was correct. This contextual paucity was aimed at bringing the experimental conditions as close to the typical q-spreading inducing contexts as possible.

Materials: target sentences with a negative determiner presented in one of the following contexts:

(25) salient extra object: No boy is hugging a dog.



(26) non-salient (multiple) extra objects: No girl is holding a cat.

(27) controls: No boy is hugging a dog/ No boy is hugging a cat.

4.4.2 Results

The children were divided into the following two groups:

- The adult-like group, who correctly rejected all but the control sentences - 6 children (and all of the adult controls).
- The 'spreading' group - 10 children, who judged as true at least three out of 4 items of at least one type. For convenience, I will refer to these non-adult-like responses 'spreading' responses.

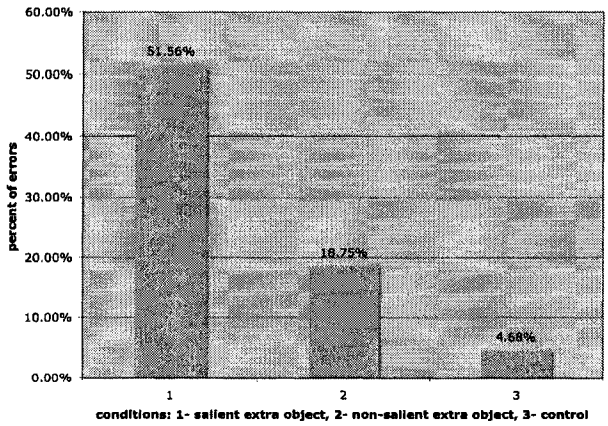


fig. 1 Error rate across conditions

Similar to Philip's 'symmetry' children, children who gave "yes" responses often explained their judgment by pointing to the extra individual in the picture and saying "not this one" or "he means this one".

Even though 'spreading' responses were given in both test conditions, the 'salient extra object condition' produced the highest rate of errors (see fig.1). The 'yes' responses were given even though the children had successfully pointed out the multiple dog-hugging boys before the test sentence was presented.

4.4.3 Discussion

We can ask whether the adverbial theory of universal quantification (Philip 1995) can be extended to account for these errors. The answer seems to be that we cannot, at least not without drastically modifying the lexical semantics of "no". If we simply adopt the event quantification semantics for sentences with the negative determiner, we would get the semantic form like that in (28):

(28) ∀e [a boy or a dog participate in e] [it's not the case that boy hugs dog in e]
"For every event in which a boy or dog (or both) is a participant, it is not the case that a boy hugs a dog".

The sentence construed this way would be falsified by any dog-hugging boy, which leaves the 'yes' responses we elicited unexplained. The correct truth conditions can, however, be captured by means of a contextually narrowed quantifier domain without altering the regular adult semantics of the negative determiner. For those children who accepted the target sentence, it is interpreted as containing an implicit domain restriction, something like the one given in (29) with its meaning showing in (30):

(29) No boy is hugging a *fluffy white* dog *over here.*

(30) ∀x [boy(x) → ¬∃y[dog(y) and y is in C and x is hugging y]]
C={x: x is a fluffy white dog}

A conclusion that we can make is that the visual asymmetry is responsible for the negative determiner errors, particularly when the asymmetric individual is highly salient. The visual context serves as a pragmatic clue, particularly when the experimental design offers no linguistic context to supply that information. This asymmetry is interpreted as an indication of the relative relevance of the individuals in the picture: the salient extra dog is perceived as the most relevant and included into the singleton domain restriction for the indefinite. This allows for a yes-response, since there clearly is a dog such that no boy is hugging it.

4.5. Finding values for the domain restriction variables in adults and children

If children's semantics of the negative determiners is adult-like, we can ask why adults don't make the same type of errors. The answer has to do with the observation that for

adults, the sentence "no boy is hugging a dog" uttered in the context of the picture with boys and dogs is most readily understood as talking about all of the individuals in the picture. As was observed by Kratzer 2004 in her discussion on covert domain restrictions in adult grammar, the utterance in (31) is most readily understood as talking about everybody in the room where the sentence is uttered. It would be judged false if only a subset of the people present in the room is smiling, even if there is a highly salient feature distinguishing the group of smilers (e.g. everyone who is smiling is wearing a red hat). Likewise, (32) would be judged false if there are any smilers in the room, even if the non-smilers share a highly salient feature.

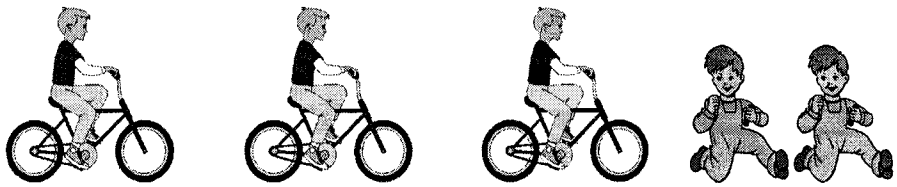
(31) Everybody$_C$ is smiling.
(32) Nobody$_C$ is smiling.

These examples show that for adults, properties of individuals, even very salient ones, are not readily picked up as values for domain restriction variables. This is, however, exactly what I claimed occurred with children in experiment 3. The contrast between adults and children becomes less drastic if we consider the following observation: the difficulty of picking up sub-properties as values for covert domain restrictions holds true only in an out-of the-blue context. Thus, if one utters 'every boy is smiling', and points at a picture, the listener would obligatorily interpret 'every' as ranging over all of the boys in the picture, and not any subset thereof, regardless of any visually distinctive features some of them may share. However, such narrowing is possible in natural discourse. Let's imagine the example given in (34) uttered in the following situation:

(33) **Situation**: department meeting where both students and faculty are present. The head of the department announces that according to the new regulations, the students who miss generals' deadlines will no longer receive department funding.
(34) A person present in the room: Now everybody is worried.

There is the intuition that the sentence may be judged true even if only the students present at the meeting are worried. From this, we see that in a natural discourse, there are contextual clues when domain restrictions are narrowed. Without overt clues, adults assume that the maximal set of individuals present in the situation has to be filled as the value for the domain restrictions. Children, on the other hand, do not always obey this strategy and may adjust domain restrictions without overt linguistic context. This may be, as I stipulated earlier, due to their over-interpreting the visual information as pragmatic clues that can be read by the interlocutors. One additional piece of evidence that this unexpected domain narrowing occurs comes from the so-called under-exhaustive errors, another type of errors usually included under the umbrella term of quantifier spreading. The error is illustrated in (35):

(35) Experimenter: Is every boy riding a bike?
Child: Yes.



While it is difficult to explain this type of response under the event quantification theory, it can be straightforwardly analyzed as the child's having narrowed the domain for 'every' to include only the boys deemed relevant, something an adult would do only if the context contained explicit clues for such narrowing.

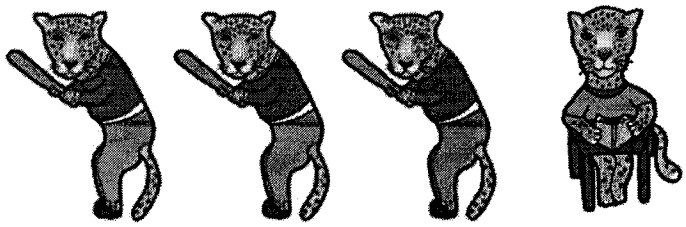
I would like to suggest that the main difference between adults and children that is responsible for "q-spreading", is not in the formal semantics of quantification, but in the

freedom children have when making assumptions about narrowing domain restrictions in situations of contextual paucity.

This phenomenon may be related to children's weaknesses in the area of the Theory of Mind, which I understand broadly as a cognitive module responsible for our making fairly accurate conjectures about the content of others' minds, including gauging how much is known to others in a given situation. I would like to suggest that the underlying issue responsible for children's mistakes in constructing quantifier domain restrictions is their weakness in attending to others' access to information and making a connection between the sources of information and their epistemic effects on others, well-documented in the Theory of Mind literature (e.g. Wimmer, Hogrefe, & Perner 1988, Taylor 1988 among others). Given this weakness, it is not unreasonable to expect children to be freer than adults in narrowing the domain restriction without overt linguistic context.

I have found additional evidence that children's domain restrictions are more flexible than those of adults. In experiment IV, which I will discuss in more detail in chapter 5, when children were shown pictures like those in (36) and asked to judge the truth of the sentence below, they unanimously accepted it as true:

(36) Puppet: One jaguar is reading a book, but every jaguar is playing baseball.
Child: True (78% in the 15 preschool and kindergarten aged children tested).



In contrast, their responses to the following controls were nearly 100% corect. The difference is that for 'both', the size of the quantifier domain restriction is fixed by lexical semantics, and "both" requires it to consist of exactly two individuals:

(37) Puppet: One cat is eating, but both cats are sleeping.
Child: No, only one is sleeping.


Adults responded to (36) with squeamishness equal to that with which they responded to (37), judging both as contradictions.

4.5.1 Experiment 3, Part II: the extra subject condition

In addition to the general 'flexibility' of quantifier domains, there was another non-adult-like property I have uncovered in experiment 3, namely in the 'extra subject' condition. In this condition, the target sentences had the same structure as in the two conditions discussed above, but the pictures contained an extra individual of the same sort as the agent, illustrated in 38:

(38) Puppet: No boy is hugging a dog.
Child: Yes (points to the extra boy)




In our sample, this type of response was given at the rate of 26.56 %, lower than in the 'extra dog' condition, but higher than in the non-salient condition (with multiple extra

dogs). Here again, the non-adult-like children seem to narrow the domain, but in this case it is the domain of the negative determiner that is narrowed to include a single boy, an option normally rejected by adults. This suggests another difference between adults and children: children allow singleton domain restrictions for those quantifiers that disallow them in adult gramamar. Thus, for adults, sentences with the negative determiner denote empty intersections, in our example between a set of boys and a set of dog-huggers, and require the set of boys to be non-singleton. The sentence is odd in a situation in which there is only one boy as a denial that the one contextually present boy hugged a dog (it is acceptable if understood as 'non-specific', denying that any boy huged a dog). Children do not seem to have this 'anti-uniqueness' requirement.

There is independent evidence that children do not obey the 'anti-uniqueness' for aanother quantifier that requires it for adults, namely the universal quantifier 'every'. In the study by Yatsushiro and Sauerland 2004, children were shown to allow singleton domain restrictions for "every", an option disallowed by adults. In their study, Y&S investigated children's knowledge of the presuppositions of 'every' – the existence and the anti-uniquness presuppositions. The former requires the context for a universally-quantified sentence to contain a non-null set of individuals of the sort denoted by the common noun of the universally quantified NP. The latter requires this set not to have the cardinality of one. These properties of the universal quantifier are illustrated in the following example. Both (40) and (41) are true in the context of (39), but each violates one of its presuppositions:

(39) Context: a boy named Johnny, his parents, his one red-headed sister, two brothers, no grandparents.

(40) Johnny: Every grandma of mine is wearing red pants. (existence presupposition is violated).
(41) Johnny: Every sister of mine has red hair. (anti-uniqueness presupposition is violated).

Y&S found that while children had a high rate of correct responses with (40) - 72% correct, with (41) there was only a low 32% of correct responses. Based on these results, as well as the results we obtained on the 'extra boy' condition, we can conclude that children allow "every" and "no" to combine with singleton domains, which are disallowed by adults. This conclusion is consistent with the Yatsushiro and Sauerland's hypothesis, which is based on Heim's theory of presuppositions. It holds that some presuppositions are inherent in the lexical semantics of the expressions, while others are implicated. The two types of presuppositions follow separate acquisition paths, with the implicated presuppositions being harder. Thus, the existence presupposition, which children know, is lexically encoded for "every", while the anti-uniqueness is derived as an implicature via Heim's "Maximize Presupposition". It states that if there are logically equivalent expressions that differ only with respect to their presuppositions, the one with the strongest presuppositions that are satisfied must be used. Consequently, using the expression with a weaker presupposition entails that the stronger presuppositions are not satisfied. According to this, the anti-uniqueness of "every" (and "no") arises as an implicature because a potentially more informative alternative ("the") is not used. Children do not reliably calculate implicatures and hence lack the anti-uniqueness for "every" and "no".

5.6 Conclusion

In this chapter I have reported the findings of a series of experiments aimed at testing certain predictions made by the theory of "q-spreading" that I developed in the chapter 3. I have claimed that children possess full grammatical competence (syntactic and semantic) of universal quantification but have difficulties with some of the pragmatic aspects of quantification, namely finding correct values for contextually determined domain restrictions. The results of the experiments reported here support the full competence hypothesis by showing that:

- Children allow both the surface and the inverse scope readings of universally quantified sentences (experiment I).
- Children have correct lexical semantics of the universal quantifier since in the surface scope condition they unanimously judged the target sentences true despite contextual asymmetry (experiment I).
- Children know the asymmetric entailment relations between the indefinite-wide scope and the universal-wide scope readings of universally quantified sentences since they make significantly fewer errors judging the truth-value of such sentences in the context of a picture consistent with the indefinite wide scope, which entails the universal wide scope reading compared with the context consistent only with the universal wide scope reading, even though both pictures contained an extra object-set individual and under the event quantification theory were expected to result in a similar error rate (experiment II).

Another conclusion that follows form the reported results has to do with investigating the nature of non-adult-like responses reported in 'spreading' literature. I hypothesized that errors reported in the literature stem from children's weak ability to cope with contextual paucity. In those experiments in which linguistic context is not given, children use visual salience to help them decide which individuals are relevant. Those 'extra' individuals that are less visually salient (are included in a set of multiple other individuals of the

same type) are interpreted as not being relevant giving a much higher rate of adult-like responses. Our results indicate that:

- children's rate of errors was higher in those experimental conditions in which the 'extra' individual was highly perceptually salient (experiment II and III).

Another hypothesis suported by my findings was that 'spreading-like' errors do not stem primarily from children's difficulties with the universal quantifier, but the indefinite, and therefore can be elicited with target sentences containing the negative determiner instead of the universal quantifier. I have concluded that the errors are caused by the following properties in which children are different from adults:

i. allowing the value of the covert domain restrictions to be set with subsets of the given individuals with a salient feature, even if the context is too poor for such domain narrowing for adults (experiment III- 'extra object' conditions, experiment IV)
ii. allowing singleton domains for the quantifiers that in adult grammar do not allow them (experiment III- 'extra subject' condition).

I have conjectured that the root cause of these differences lies in children's weak TOM-related abilities to determine what is known and what is relevant to others. Further research is needed to test the correlation between q-spreading and the weakness in knowledge attribution. I believe that this presents a productive avenue for further research and can lead to a better understanding of other types of interpretive errors in both children and various populations of adults currently being attributed to non-adult-like semantics.

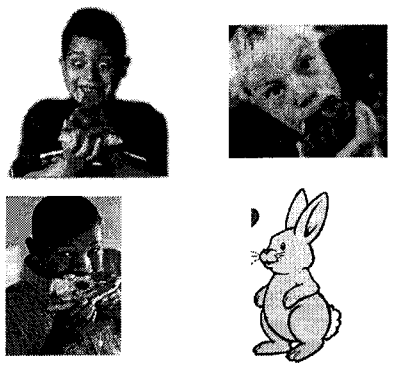
Chapter 5 Perfectionist errors

1. Perfectionist responses: overview

The main goal in this chapter is to understand the nature of the remaining type of "q-spreading" errors – the so-called "perfectionist" and "bunny-spreading" errors, i.e. to explain why some children reject examples like (1) in picture-contexts like (2) and (3):

(1) Every boy is eating pizza.

(2)



(3)

Philip (1995) discovered that for a group of children, sentences like (1) in contexts analogous to (2), where, in addition to a number of pizza-eating boys, there is another individual, in this case a bunny, with or without pizza, resulted in a relatively high rate of error (53%). He called such responses and the children who produced them "perfectionist". Roeper et al. 2004 found a similar error in contexts like (3), where the unmentioned individual is not simply present or engaged in the same activity (eating

pizza), but is doing something else (e.g. eating a carrot), a response they termed "bunny-spreading". They proposed a maturational account according to which, such errors occur at a developmental stage when the quantificational determiner *every* is interpreted as an adverbial element (roughly synonymous with *always).* According to Roeper et al., at this "bunny-spreading" stage, the child does not project a DP above the NP and instead interprets subject quantifiers as adverbial elements in Focus Phrase, a projection dominating CP. Consequently, the quantifier takes scope over the entire sentence, quantifying over events in a manner similar to Philip's analysis, and a presence of any individual other than a pizza-eating boy would falsify the sentence for the child at this stage of development.

The "perfectionist" and "bunny-spreading" errors appear to be very similar to each other, and both indicate a response when a universally quantified sentence is rejected by the child when the picture-context contains an extra, unmentioned individual of equal visual salience as members of the agent group. It is not clear from Roeper et al.'s discussion whether they analyze "bunny-spreading" as a distinct type of error from Philip's perfectionist errors. However, because of their close similarity, I will use the term 'perfectionist' to refer to both of these types of errors. I will refer to the types of context in which the errors arise as the 'unmentioned individual condition'. As reviewed in chapter 2, the "perfectionist" semantics posited by Philip for sentences like (1) requires for every sub-event of the event under consideration that involves **any participant of any sort** to satisfy the truth conditions of the nuclear scope, as shown in (4):

(4) ∀e [C(e) & any individual participates in e] [boy eats pizza in e]

According to this unorthodox semantics, the set of sub-events e forming the restrictor of the universal quantifier are the sub-events of the contextually relevant event that meet a particularly broad restriction, namely that any individual participates in *e*. The sentence is then falsified by any sub-event *e* with any participant that doesn't satisfy the truth-conditions of the nuclear scope, namely that a boy eats pizza in *e*.

This analysis captures the negative responses to (1) that some children may give when the picture-scenario contains a participant that is not a pizza-eating boy. It, however, lacks an explanation for why the initial developmental stage should correspond to the "perfectionist" semantics. The maturational hypothesis by Roeper et al. maintaining that children lack the DP projection and generate the universal quantifier in FP is insufficient since by itself it does not capture the "perfectionist" semantics. Thus, the "perfectionist" interpretation in (4) is not equivalent to the adult interpretation of sentences with quantificational adverbs like 'it's always the case that the boys eat pizza', which, according Roeper et al. 2004, it should be roughly synonymous with. For adults, this sentence is not falsified by unmentioned participants eating pizza or doing something else, such as sleeping, dancing, etc., but is falsified by some or all of the boys failing to eat pizza.

The proponents of the adverbial theory of universal quantification appeal to the mechanism of adverbial quantification in adult grammar to ague that since it exists in adult languages, it is 'natural' to posit a stage in child language when all (universal) quantification is adverbial. However, this argument is not entirely convincing. What is unusual for the adult grammar is not the notion of a quantifier being covertly moved to FP or it quantifying over events, but the type of the restrictor that the "perfectionist"

semantics has to posit.[1] If there is an account of perfectionist responses that does not have to resort to a grammatical explanation for 'perfectionism', it would have an advantage of not needing an ad hoc semantics like the one in (4).

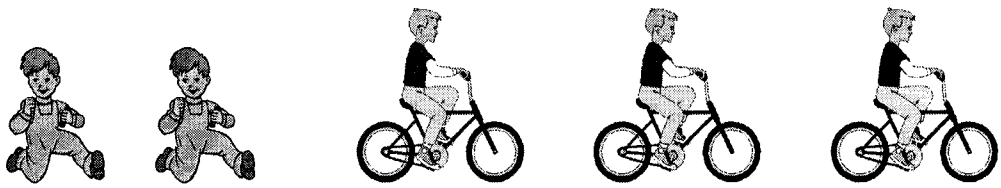
There is another issue that I believe should be addressed. Each of the different types of q-spreading has been reported to follow a distinct pattern. At the same time, intuitively, they all seem to be related. Even though it is possible that the various types of the reported q-spreading responses (classic spreading, under-exhaustive and perfectionist responses) are due to unrelated causes, a theory able to account for each type of error and its distinct pattern and as the same time capture their apparent similarity would have an advantage over those theories that treat each type of "q-spreading" as a separate phenomenon.

I will address the question of how "perfectionist" responses fit into a general picture of q-spreading. I will propose that perfectionist errors are the clearest indication that the root of the problem in "quantifier spreading" lies outside of grammar and in the pragmatic conditions of the test. I will argue that all types of "q-spreading" stem from a common property of children' cognition, namely their weakness in dealing with other minds, which I employed in my explanation of "classic spreading" errors. As I have argued in chapter 3, one way in which the developing Theory of Mind may interfere in children's performance on certain types of tests is to make them choose the unintended scope reading of a scopally ambiguous sentence. One instance of this is when children select a singleton domain restriction for the indefinite when the picture-scenario contains

[1] The problem of a nonstandard restrictor also applies to the adverbial account of classic spreading. It casts doubt on the claim by the proponents of the adverbial theory of universal quantification that since adverbial quantification is part of adult grammar, the 'adverbial' stage of universal quantification does not constitute a non-UG option.

a salient single individual to draw the child's attention and cause her to reject a universally quantified sentence in the "extra object" condition. Another manifestation of this is when children narrow the domain of the universal quantifier to include only the 'relevant' individuals, producing the so-called under-exhaustive errors:

(5) Experimenter: Is every boy riding a bike?
Child: Yes (points to the three boys riding bikes).



Another type of 'miscommunication' that may occur has to do with the "Question under Discussion', the question that the target sentence is assumed by the child to be answering. If the experimenter does not pose an explicit question for which a universally quantified proposition becomes a pragmatically acceptable answer, the child may fail to posit such a question and as a result judge the target sentence unacceptable, the phenomenon that I will propose explains the "perfectionist' responses.

2. The pattern of the two types of 'q-spreading' (Philip 1995)

In order to understand the general pattern of q-spreading better, I will re-examine the pattern of responses reported by Philip (1995), one of the most extensive studies on this topic to date. In this study, Philip identified three response patterns with respect to the interpretation of (1). One group of children in his sample was adult-like. It was the smallest in number (n=32) and did not produce "spreading" errors in any of the experimental conditions. Hence, nothing needs to be said about them. The two other

groups differed from each other not only in the types of errors they produced, but also in the rate of errors.

One of the two remaining groups was the "symmetry group" (N=87), which included those children who produced non-adult-like responses in the "extra object" condition and gave exclusively adult-like responses in other conditions. An interesting observation about this group is that their non-adult-like responses (what I have been previously referring to as "classic spreading") were produced at the rate of 56%, a fairly close split between adult-like (yes-) and non-adult-like (no-) responses. Although such split may be interpreted as at-chance performance indicative of children's deficient semantic competence, there is another explanation that captures this pattern of responses well. I have argued that the errors are connected to ways in which scopal ambiguities in these sentences may be resolved. In this view, the split between the yes- and no-responses indicates children's oscillating between the two available interpretations of the target sentence (both compatible with the given visual context). The non-adult-like responses here are pragmatic in nature, perhaps due to children's not obeying the charity principle (which should lead them to select the interpretation under which the sentence is true in a given context) and instead selecting the reading based on the visual salience of the 'extra' object (perceived some of the time to be the intended candidate for the singleton domain restriction). I propose that Philip's 'symmetry' children are those children who allow both the 'multiple' and 'singleton' interpretation for the indefinite and are not strongly biased towards either interpretation for the sentences that have the universal in the subject position (for which adults are typically biased towards 'multiple reference' or distributive interpretation). The same group is presumably unbiased

towards either interpretation for the object-universal sentences (in which adults typically have a single-reference bias), i.e. they constitute the same group that accepted the inverse scope interpretation for such sentences in experiment 2 described in the previous chapter. In a picture-reading task, this lack of a bias would allow the child to be easily led towards an alternative reading, particularly the singleton indefinite interpretation for (3), especially if the situation contains a perceptually salient single individual to be a candidate for the single reference. Predictably, in the conditions with some other type of contextual 'asymmetry', (e.g. a condition in which in addition to the donkey-riding farmers mentioned in the sentence, the picture contains a bicycle-riding bunny instead of an extra donkey), this group responds like adults - in the absence of a salient single donkey, the 'singleton' interpretation does not arise.

This analysis does not only explain Philip's results, but also straightforwardly accounts for the findings reporting a significant drop in the rate of errors brought by enriching the linguistic context (and thus clarifying the intended domain restrictions). Thus, Crain et al. (1996) found that when the experimental design is modified in such a way that the Condition of Plausible Dissent is satisfied (concurrently with providing rich linguistic context), the rate of children's correct responses in the 'classic spreading' context reaches a high 88%. Furthermore, studies like Drozd & van Loosbroek 2001 (cf. also Drozd 1998, Musolino 1998, Brinkmann et al. 1996) showed that a similar drop in spreading errors can be achieved by various other means of enriching context that do not involve Plausible Dissent. All of these findings are compatible with the explanation I have proposed.

In another group of findings, Sugisaki and Isobe 2001 showed that with an experimental design similar to the one that typically elicits a high percentage of spreading responses, the errors can be dramatically reduced by simply changing the number of extra individuals in the object set from one to multiple. S & I tested two groups of Japanese-speaking children using methodology similar to Philip's. Group 1 was presented with pictures with one extra object (three cats each bouncing a ball plus one extra ball), Group 2 with pictures with at least 4 extra objects (three cat-ball pairs plus several extra balls). They report that Group 1 had a rate of errors of only 37.5% correct, while Group 2 a high 87.5%. This result cannot be accounted for by the adverbial quantification theories, which claim that children require an exhaustive pairing of subject- and object-set individuals, since this requirement would be violated regardless of whether the number of extra objects is one or more. Likewise, the Plausible Dissent Hypothesis would be inadequate to explain these results since in both conditions Plausible Dissent was not met. Philip and Lynch 1999 achieved analogous results by manipulating fore- vs. backgrounding of the extra objects in the picture. These are again compatible with my analysis, which connects the errors with visual salience of the 'extra object', reducing which leads to a significant drop in error rate.

Now let's consider the third group of children reported by Philip, which presents an interesting puzzle. They were the so-called 'perfectionist' children (N=97), who rejected target sentences like 'every farmer is riding a donkey' not only in the "extra donkey" (hence forth 'extra object') condition, but also in the unmentioned individual (either shown as riding a donkey or not) condition. Interestingly, these children not only had more error types, but they also had a much higher rate of errors in the 'extra-object'

condition compared to the "symmetry children', namely over 80% for the former compared with the slightly over 50% for the latter. Philip concluded that even though both 'symmetry' and 'perfectionist' children made errors in the 'extra object' condition, their errors in this condition should be analyzed with different mechanisms of interpreting universally quantified sentences, as reviewed in chapter 2.

The following table summarizes Philip's results by showing the percentage of non-adult-like responses given by the 'symmetry' and 'perfectionist' groups.

Group / Picture Type	Classic spreaders (n=87)	Perfectionist (n=97)
"extra object"	57%	85%
"unmentioned individual"	-	53%

Table 1

How can we explain the finding that the group of children who made perfectionist errors also made errors of the classic spreading type at a much higher rate? Likewise, why was their rate of error across the two types of context different? Philip's "perfectionist" semantics can only handle the additional error types but provides no explanation for their different rates. I would like to provide an analysis for the 57%-85% contrast in the rate of error across the two groups, as well as the 85%-53% contrast within the perfectionist group.

3. Proposal

The first question I will address is why children make perfectionist errors. I agree with the idea that a different pattern of responses given by 'perfectionist' children represents a group distinct from the 'symmetry' children. I, however, disagree with the idea that their "perfectionism' is semantic and lies in their restricting the universal quantifier with "those events *e* that have any perceived individual participating in *e*". I will maintain that

just as it is for adult-like and 'symmetry' children, for perfectionist children 'every' is restricted by its sister NP. I propose that their 'perfectionism' is not semantic, but pragmatic in nature.

I propose that such odd responses should be accounted for with the same type of consideration proposed by Hulsey et al. 2004 for children's preference for isomorphic responses on scopally ambiguous sentences. As I described previously, Hulsey et al. discussed sentences containing negation and an indefinite, as in (6), which children typically interpret as (7), and fail to give it the interpretation in (8):

(6) The Troll didn't deliver some pizzas.
(7) It is not the case that the Troll delivered some pizzas.
= The Troll didn't deliver any pizzas.
(8) There are some pizzas that the Troll didn't deliver.

Hulsey et al. suggested that children typically respond to such sentences isomorphically (choose the surface scope interpretation) because of a pragmatic principle they called the Question-Answer Requirement (QAR). According to the QAR, children must interpret the target sentence they evaluate as an answer to a question that the discourse makes salient. Hulsey et al. propose that in all of the cases in which children fail to select the inverse scope when it is available for adults, the discourse makes a particular question salient (sometimes unintended by the experimenter). Furthermore, in all of these cases, the sentence evaluated by the child constitutes the most pragmatically natural answer to this question under its surface scope interpretation. Changing the experimental paradigm by making the target sentence an answer to the salient question under its inverse scope interpretation leads children to accept the inverse scope.

According to this theory, differences in scope assignment for scopally ambiguous sentences between adults and children do not lie in a difference in their grammatical competence (see also Musolino and Lidz 2003). Rather, adults and children differ in their

ability to accommodate a Question under Discussion. Thus, when sentence in (6) is given in a context of a story in which Grover orders four pizzas from the Troll, and the Troll is supposed to deliver all four of them, but is driving too fast and loses two on the way, when children are asked to judged an answer to the question "did the troll deliver all of the pizzas', children accept (6) as true and hence demonstrate that they accept inverse scope. In tests, when such question is not posed directly but the pizza delivery story is simply acted out before presenting the target sentence like that in (5), children often reject it and point out the pizzas that were delivered successfully as an explanation for their rejection. In this case, the QUD they are responding to is a general one like "what did Grover do" or "did Grover deliver pizza".

I believe that a similar explanation accounts for the perfectionist responses. When children are shown a picture containing a set of individuals *X* engaged in an activity, and an equally salient individual *Y*, and then asked whether or not a sentence containing a universal proposition describing the first group ('every X is ...') is true, the question intended by the experimenter and implied by the test sentence is something along the lines of "what is every X doing in the picture". However, the child is likely to be responding to the more general question "what is in the picture" or "does the puppet know what is in the picture". An acceptable answer to the former question has to be exhaustive, and, similarly, the knowledge attribution can be made only if the puppet knows an exhaustive answer to the embedded question. Thus, when the puppet utters (1) in the context like (3), it fails to mention an object present in the picture (a bunny). This frequently leads the child to reject the sentence as unsatisfactory. As an explanation for his rejection, he often points to the unmentioned individual (and says, "the bunny is eating a carrot').

This hypothesis explains why such errors occur at a significant rate only in experiments where no sufficient linguistic context is given (and consequently where no explicit QUD is put forward). Instead, the context is presented in the form of a minimally informative picture, and children have to accommodate the contextual paucity and to adjust

the QUD from the most neutral one in this situation – a general question 'what is in this picture'- to a more narrow one about the relevant set of individuals.

This hypothesis also suggests a plausible explanation for the difference between adult and child responses in typical "bunny-spreading" contexts. The reason why adults do not make 'perfectionist errors" (even though according to anecdotal evidence they occasionally make classic spreading-like errors) is their greater pragmatic sophistication. A sophisticated test-taker would be successful in recognizing that when the experimenter asks them to judge a sentence that contains a universally quantified NP 'every X', he specifically targets the test-taker's semantic competence of this expression, and, hence, only the set of Xs is relevant for the task. After all, this construction is quite infrequent in everyday discourse, and would not normally be used for a general inquiry about the content of the picture[2]. Such pragmatic sophisticates would recognize that the question the experimenter intends to have answered by the target sentence is something like: 'what are the Xs doing in the picture' and would give a correct truth value judgment recognizing that any additional individuals included in the picture should be ignored because they are there to distract the test-taker. On the other hand, those who are less sophisticated in the art of taking psycholinguistic tests, would have few clues as to which question the target sentence is meant to be answering, information normally present in natural discourse. While it is not ruled out that the child may indeed choose the correct question, as is evidenced by the significant percentage of adult-like responses, it is in no way guaranteed. Frequently, they fail to do so and instead have to judge whether an under-exhaustive answer to the question 'what is in the picture' is acceptable. It is not surprising, then, that under these circumstances children frequently reject it as unacceptable responding with a 'perfectionist' answer.

Thus, I have proposed that "perfectionist errors" stem from a flaw in the experimental design, which fails to clarify the relevant question the target sentence is

---

[2] Roeper et al. 2004 reports that the frequency of 'every' is extremely low: ...

meant to be answering. Consequently, children interpret their task as a decision about whether or not the target sentence describes the picture accurately, which means not only correctly, but also exhaustively. When they reject the target sentence in the "extra participant' conditions, they are not simply judging its truth value, but its exhaustivity as an answer to a certain QUD questioning the composition of the picture. Since it does not answer this question exhaustively, it is judged as an inappropriate answer and is rejected.

Under this proposal, the two groups – 'classic spreading' and 'perfectionist' do not represent two different stages of development. If this proposal is correct, the nature of the "perfectionist" errors is similar to the 'symmetry' error in that both groups fail in their judgment of what is relevant. The 'symmetry' children choose the correct domain restriction half of the time, while erroneously focusing on the extra item on the rest of the trials. Similarly, the 'perfectionist' children are correct in their decision about the relevant individuals about half of the time, but take everyone to be equally relevant the rest of the time; hence a similar rate of "perfectionist" errors and "classic spreading" errors for the non-perfectionist group (around 52% for the former and 57% for the latter). In the next section, I will try to clarify the role of exhausivity in perfectionist errors by briefly discussing the theory of exhaustivity and acquisition research that investigated children's knowledge of exhaustivity.

4. Exhaustivity

I will adopt a view of exhaustivity for wh-questions based on Heim 1994 and subsequent work that built on it (Beck and Rullmann, Sharvit 2002). According to this view, exhaustivity is not part of the interrogative semantics, as argued, for example, by

Groenendijk and Stokhof (1982, 1984), but a consequence of the way question meanings get integrated semantically and interact with their linguistic context. A crucial factor in producing or limiting exhaustivity for questions is the pragmatic requirement of giving not only a true, but an appropriately informative answer, which under certain circumstances requires to be maximally informative, while under other circumstances may be under-exhaustive (an example or the so-called 'mention-some' answer). What is appropriately informative depends on the specific context: it may be an answer that is weakly exhaustive, strongly exhaustive or an example answer containing a true element of the question's denotation.

The notion of strong exhaustivity was introduced by Groenendijk and Stokhof (1982, 1984). It is illustrated in the following example. Knowledge reports like that in (9a) require strong exhausivity. Sentence in 9a entails not only 9b, but also 9c. In other words, when Simon knows the answer to the embedded question (*who is in the kitchen*), the information he knows includes not only all of the individuals who are in the kitchen, but also does not include a mistaken belief about some individual being in the kitchen, who in the actual world is not in the kitchen. Thus, if Simon, in addition to believing correctly that Tom, Dick, and Harry are in the kitchen, also mistakenly believes that Susan is in the kitchen, (9a) fails to be true, even though it satisfies (9b). Under the weak exhaustivity, 9a would entail only 9b, and not 9c, a wrong prediction in this case.

(9) a. Simon knows who is in the kitchen
b. For every X who is in the kitchen, Simon knows that X is in the kitchen
c. For every X who is not in the kitchen, Simon knows that X is not in the kitchen

However, in other cases, weakly exhaustive answers are allowed. Consider the following:

(10) a. Simon knows who plays the violin.
b. For every X who plays the violin, Simon knows that X plays the violin.
c. For every X who does not play the violin, Simon knows that X does not play the violin.

In (10), (a) entails (b), but only about a subset of individuals – made relevant by the context, perhaps the set of mutual acquaintances; it does not, however, entail anything about individuals outside of the relevant set. Furthermore, it may not entail (10c) even for the set of mutual acquaintances. Let's suppose that the relevant set of individuals who can play the violin include Alice, Boris, and Carla, while Donna and Evan play the cello. Simon may correctly believe that Alice, Boris, and Carla can play the violin. He may be agnostic or mistaken about Donna and Evan, and not know that they do not play the violin. We can still attribute knowledge of the answer to the embedded question to Simon in certain situations. For example, if the question is asked in the context of the questioner looking for violinists to audition for a string quartet, rather than identifying musical skills of every member of the group. By uttering (10a) in this situation, the speaker attributes to Simon the knowledge of a weakly exhaustive answer to the embedded question, but not necessarily a strongly exhaustive one: as long as Simon's belief-worlds include individuals who play the violin in the actual world, strong exhaustivity is unnecessary and therefore is suspended. Furthermore, sometimes a question does not require even weakly exhaustive answer, but allows or even favors an example answer ("mention-some" answer); e.g. *Sue knows where you can buy a used violin.*

Heim (1994) developed semantics of exhausitivity as a property independent from the denotation of questions building upon Karttunen's (1977) analysis of embedded questions and *know*. Her approach allowed to capture maximality of answers required by some types of questions (e.g. degree questions like 'how high can John jump') without positing any separate maximality operators. According to this approach, exhaustivity is not distinct from maximality, and both are captured not by stipulating them as part of the question's denotation, but by the notion of a true and complete answer, which captures both the weak and strong exhaustivity, while allowing for the mention-some readings.

Karttunen proposed that the meaning of a question is the set of true propositions that constitute its answer. The denotation of the question (11) is given in (12) and its alternative representation (for the above scenario) in (13)[3]:

(11) Who plays the violin?
(12) $\lambda p \exists x$ [person(w)(x) & p(w) & p = λw'[x plays the violin in w']]
(13) [[who plays the violin]]^{w}={p$\in$ D <s,t>: $\exists$x[x$\in${A, B, C} & p(w) & p= that x plays the violin]}= {Alice plays the violin, Boris plays the violin, Carla plays the violin}

Karttunen's representation of 'know' looks like (13):

(13) [[know]](w)(q)(x)=1 iff
(i) x believes $\cap$q(w) in w, and
(ii) if q(w)=$\varnothing$, then x believes λw'[q(w') = $\varnothing$] in w.

Clause (i) states that x believes the intersection of the sets constituting questions extension q(w), which for an embedded wh-question amounts to a conjunction of the true propositions from the question's extension in w. Clause (ii) was added to account for the intuition that if the set of those who play the violin is empty, to know who plays the violin entails knowing that the set is empty. Without this added stipulation, clause (i) would falsely predict that if the set of violin players is empty, and hence the intersection of the set contains no propositions, *Simon knows who plays the violin* would be always vacuously true, regardless of the content of Simon's believes in actuality.

As observed by Groenendijk and Stokhof, and subsequently Heim, Karttunen's analysis has limitations; namely it fails to account for strong exhaustivity (as well as de dicto readings). In order to capture strong exhaustivity, Heim modified clause (ii) in (13) by making it more general:

---

[3] In subsequent work building on Karttunen's original proposal, the content of the question extension was modified from containing only true propositions, as assumed by Karttunen, to containing all possible alternatives (both true and false) from which the addressee is to choose his answer (in the style of Hamblin 1973). According to this, the question denotation in (12) should be modified as follows:
$\lambda p \exists x$ [person(w)(x) & p = λw'[x plays the violin in w']]
I am going to put aside precise details of alternative approaches to interrogative semantics since they are of no import for the question at hand.

(14) [[know]](w)(q)(x) = 1 iff
(i) x believes ∩q(w) in w, and
(ii) x believes λw'[q(w') = q(w)] in w

This "generalized Karttunen analysis" captures the intuition that knowledge of the extension of an embedded question can only be attributed if the subject not only knows the conjunction of all true propositions constituting the answer to the question, but also knows that false alternatives are not part of the answer. In Heim's words, the subject "knows that the answer is the actual answer". As Heim goes to show, in this formulation clause (i) is redundant because it follows form (ii), and hence, can be dispensed with in the lexical semantics of the question-embedding verbs like *know*.

The distinction between (i) and (ii), which capture the difference between weak and strong exhaustivity is preserved as separate semantic objects, namely in Heim's notions of asnwer1 and answer2, the two notions that she argued are needed to define a true, appropriate answer. Answer1 is true and complete answer and denotes the intersection of the question's extension. Answer2 is a strongly exhaustive answer - a proposition that answer1 to question Q is what it is in the actual world. Thus, according to this approach, interrogative semantics includes the following semantic objects: the basic denotation of a question (Kartunnen style set of propositions), answer1 and answer2.

(13) Answer1(w)(Q) = ∩ { p: Q(w)(p) & p(w) } – the conjunction of all true propositions in the question extension; i.e. the intersection of all true members of Q(w).

(14) Answer2(w)(Q) = λw' [answer1(w')(Q) = answer1(w)(Q)] – the proposition that the answer1 to Q is what it is in the actual world w.

Answer1 not only insures that no false propositions are included in the resulting set, but by forming an intersection (i.e. a conjunction) of the true proposition insures the maximality of the answer (its completeness). Thus, the embedded question in (10) under

weak exhaustivity as in (10b), would allow the set given in (13) to be the complete answer (Answer1).

Beck and Rullmann (1999) discussed degree questions; e.g. *how high can John jump,* which requires a maximal degree answer, or *how long does one have to work to get a paid vacation,* which requires a minimal degree answer. Their analysis of such data provided evidence that the notion of Answer1 is able to capture what counts as a complete answer to a degree question without using any additional devises of maximality or minimal sufficiency. They also showed that another type of answer is needed in order to capture strong exhaustivity, and that Heim's Answer2 works well for this purpose. Thus, for a non-scalar question like (10), it adds to Answer1the information that nobody but Alice, Boris and Carla plays the violin. For a scalar question it adds the information that the answer is maximal.

The issue especially relevant to my goal is B&R's flexible approach to how the choice between strong and weak exhaustivity or "mention some" interpretation is determined (at least for those embedding verbs that allow their interrogative complements to have either).[4] They proposed a rich and flexible system, which provides questions with various degrees of exhaustivity, a system in which the choice of answer is determined by the factors external to the question itself (i.e. by an interplay between the semantic properties of the question-embedding predicate and the context). The role of pragmatics is especially crucial in determining the degree of exhaustivity in unembedded questions. It involves using Gricean maxims to determine how informative the answer should be. If the context suggests that the questioner will be satisfied with an example

---

[4] The debate about the correct account for the weakly exhaustive embedding verbs is not importsnt for the type of data I am investigating and hence

answer, an addressee will provide an answer containing only a subset of the true elements of Q(w) instead of a complete list. Under other circumstances, the addressee would determine that a complete answer is required – answer1. This in turn would trigger the implicature that the interlocutor giving an answer is being as informative as possible, and Answer2 is inferred from Answer1. Certain contexts may make strong exhaustivity unnecessary, and hence cancel the inference from answer1 to answer2.

This flexible approach to exhaustivity requires the speaker not only to know the Karttunen/Hamblin denotation of questions and the notions of Answer1 and Answer2, but have competence of pragmatic nature, being able to judge what type of answer will have a desired degree of exhaustivity in any given situation.

4.1 Exhaustivity - Perfectionism Connection

Now let's go back to the "perfectionist" responses. The errors occur when the child is instructed to judge whether the puppet 'knows what is in the picture' without making it explicit that the puppet is not required to 'mention everything' and is allowed to give a 'mention-some' answer, (or if the child is asked to judge whether "the sentence goes with the picture" and would most likely evaluate whether the target sentence provides an acceptable answer to the question "what is in the picture". Under these circumstances, the child is likely to assume that the target sentence containing a universal proposition about a subset of individuals in the picture is intended to be strongly exhaustive (to infer Heim's Answer2). Under these circumstances, by uttering 'every dog is eating a bone', the puppet is perceived to be making an additional implication that there is nothing else in the picture. This is plainly false given the presence of a rabbit eating a carrot.

(15) Question under Discussion: What is in the picture?
(16) Exhaustivity options:
"mention some": e.g. dogs eating bones;
"weakly exhaustive": dogs eating bones and a bunny eating a carrot;

"strongly exhaustive": dogs eating bones and a bunny eating a carrot, and nothing else.

Answering with the "mention some" option is either not informative enough (if weak exhaustivity was expected) or false (if strong exhaustivity was assumed).

4.2. Exhaustivity Experiment (experiment 4)

4.2.1. Goal, Participants and Method

Goal: to test the hypothesis that the perfectionist errors are a child's negative evaluation of the target sentence as an appropriate answer the question "what is in the picture".

According to a grammatical account of perfectionist responses, children go through a semantic stage in which the universal quantifier "every" functions as an adverb and its restrictor is formed by all sub-events of the relevant event that have any participant. In contrast, according to my hypothesis, they produce the perfectionist responses because they expect the target sentence to be an exhaustive answer to the question "what is in the picture". To test this hypothesis, the experimental design can be modified in such a way that the exhaustivity requirement is made explicit and the puppet's response is exhaustive. Consider the following example:

(17) Experimenter: Does the puppet know what is in the picture?
(18) Puppet: Every boy is eating pizza and a bunny is eating a carrot.
(19) Puppet: Every boy is eating pizza.

(20)

According to our null hypothesis, the universally quantified sentence in both (18) and (19) should be false for the perfectionist child in context (20) since in both cases the presence of another participant (whether it is unmentioned as in (18) or mentioned as in (19)). Thus, if the grammatical stage account is correct, mention-all responses should not significantly reduce the error rate. However, my hypothesis predicts that perfectionist errors would not occur with (18).

Participants: 15 preschool and kindergarten children attending UConn Child Labs (median age = 4;8); adult controls (N=5)

Procedure: a version of the standard Truth Value Judgment task. The context was presented in the form of a picture and the child was asked to judge whether the puppet knew what was in the picture by evaluating the target sentences describing the picture "uttered" by the puppet. The task included an explicit reminder: "make sure he (the puppet) tells us about everything and does not leave anyone one."

Materials: 4 universally quantified sentences like (18) presented in each of the following contexts (20 test items total plus fillers):

(21) "Exhaustive" conditions:

a) 'perfectionist' (the same as in (20));
b) "extra object":

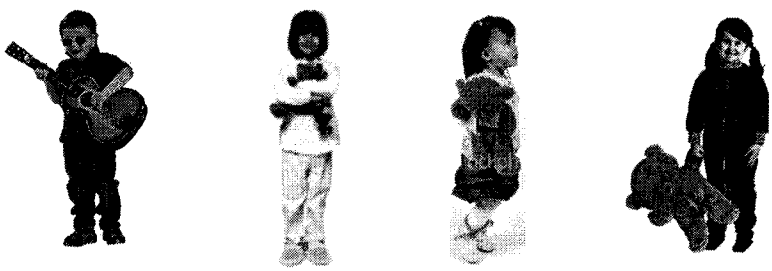
Every dog is wearing a hat and there is another hat.

(22) "Under-exhaustive" conditions:

a) Type 1: Every girl is reading a book.

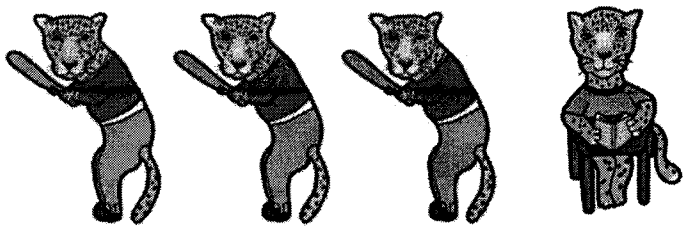
b) Type 2:

Every girl is holding a teddy bear.



(c) "contradictory":

One jaguar is reading a book, but every jaguar is playing baseball.



In both "exhaustive" conditions a and b, the children were predicted to have a comparable high rate of correct "yes" responses across both conditions. In the under-exhaustive condition, the puppet fails to mention the 'extra' individual: in the under-exhaustive condition of type 1 an individual of the same sort as the agent; in the under-exhaustive condition of type 2 of another sort. Both conditions were expected to generate correct "no" responses from children due to under-exhaustivity. The "contradictory" condition was added as a control. Children were expected to reject these sentences as containing a contradiction: the first and the second conjunct cannot both be true.

The idea of using this type of control sentences came up in connection with the null hypothesis about children passing through the "perfectionist" and "classic spreading" semantics stages. If this hypothesis is correct, then stimuli (18) and (21) should be analogous to the contradictory (23). According to the event quantification semantics, for the sentence "every dog is wearing a hat" to be true, every sub-event in which either a dog or a hat participates (or in which anyone participates for the perfectionist stage) should be an event of a dog wearing a hat. Therefore, an extra hat in case of (21) or a bunny eating a carrot in (18) falsifies the first conjunct and therefore both conjuncts cannot be true. Thus, the assumption was that if children accept these sentences as true, their semantics must be run-of-the-mill and not 'event quantification'.

4.2.2. Results

Our main predictions were confirmed. The results indicated that in the two "exhaustive" conditions children's responses contained a very high percent of correct responses. In the whole data sample in these two conditions (120 responses across 15 children) there were only 9 instances of incorrect "yes" responses (or 7.5%). The 'under-exhaustive'

conditions resulted in a higher error rate – 26 out of 120 instances of incorrect acceptances (or 22%), which is still significantly lower than the 53% of under-exhaustive errors obtained by Philip 1995, where there was no explicit mention of exhaustivity (or rather a reminder that exhaustivity was not expected).

The most surprising result was the children's responses in what I called the "contradictory" condition. Contrary to my expectations, children overwhelmingly accepted these sentences at the rate of 78%. In the sample, only 3 children correctly rejected at least 3 of the 4 sentences of this type. The results are summarized below in figure 2.

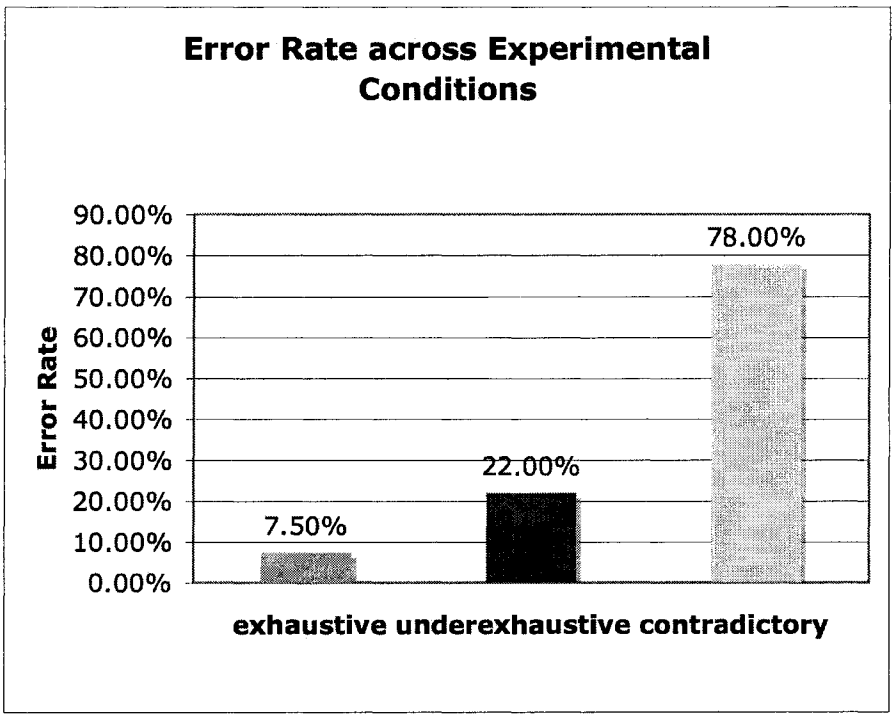


Figure 2

Discussion. The results of this experiment show that when the exhaustivity requirement is made explicit, children's responses become predictable: children correctly accept those

universally quantified sentences that are part of an exhaustive answer about the content of the picture scenario and reject those that constitute an under-exhaustive answer. This supports the hypothesis (even if it doesn't directly prove it) that the perfectionist errors reported in the literature are due to the under-exhaustivity of the test items and a failure on the part of the experimenter to clarify that exhaustivity is not required. A more direct test of this theory, which I will leave for further research, would be to conduct a similar experiment, but to tell the child that puppet's statements do not need to be exhaustive and that the puppet should be rewarded if he makes any correct statement about the picture regardless of whether it is a complete description. If my hypothesis is correct, such experiment should produce a high rate of correct responses, close to the rate I have obtained in this experiment in the "exhaustive" condition.

An interesting finding comes from what I called the "contradictory" condition, which children were expected to reject (as was the case with the 5 adult controls). However, this is not what I found. Instead, children overwhelmingly responded by judging these sentences acceptable. In chapter 3, I have discussed the evidence that children's domain restrictions are more flexible than those of adults; i.e. they allow narrowing the domains under the circumstances that adults do not. This seems to be the most straightforward explanation for these responses: children do not see sentences like (23) as a contradiction because they are able to narrow the domain of "every" to include only the relevant jaguars (those who are playing baseball). If this is the case, my expectations that children with "the event quantification" semantics should reject the exhaustive sentences (like (18)) as either false or a contrdiction may not be warranted and children may resort to a similar domain narrowing with event quantification, leading

them to accept these sentences, just as I found. This undermines, somewhat, the strength of the evidence I found and leaves a possibility that children in fact do possess 'perfectionist' semantics. However, even though both my proposal and the event quantification theory can account for the improved results I have reported by appealing to domain narrowing, the more parsimonious account is the one who needs no additional mechanism to do so. In this respect, the event quantification theory is fares less well. In addition, there is independent evidence that exhaustivity is connected to 'q-spreading', which I will discuss in the next section.

4.3. Previous Studies on Children's Knowledge of Exhaustivity

Strauss 2002 argued that the acquisition of exhaustivity is tied to the acquisition of universal quantification, and that it is typically the same children who make both sorts of errors. He looked at the corpus of data collected by the DSLT project (Seymour et al.). In this corpus from 1295 children (normal and disordered) between the ages of 4 and 12 (median age 6; 6), among other measures, there were included 9 questions of three types: a wh-question (requiring an exhaustive answer), an echo-question and a multiple wh-question (requiring an exhaustive pair-list answer). Unfortunately, the test did not contain questions requiring a non-exhaustive answer, so we do not have data of children's over-exhaustive errors from this study. The results, however, indicate that under-exhaustive answers were quite prevalent with younger children having the highest rate of such errors and the rate of errors gradually declining with age. Overall, 39.3% of the non-disordered children made one or two errors on the wh-questions. The data reported by Strauss is summarized in (24):

(24) Under-exhaustive answers by age:

Age:	4	5	6	7	8	9	10	11	12
Errors/child:	0.93	0.48	0.38	0.37	0.15	0.15	0.02	0.02	0.09

Strauss hypothesized that there may be a connection between errors in exhaustivity and errors with universal quantification. He found that there is a correlation between the two types of errors:

(25) # of Wh-exhaustivity errors

# of Universal errors	0	1	2+
0	84.9	11.2	3.9
1	63.9	23.3	12.9
2	59.8	22.4	17.6

This table indicates that making 1 or more universal errors correlates with a higher rate of wh-exhaustivity errors. He hypothesized that if the two types of errors are independent from each other, one should expect children who make 1 or more errors with the universal quantifier to make errors of exhaustivity at the same rate as those children who did not make any universal quantification errors, the prediction summarized in (26). The actual results, however, were different and are given in (27).

(26) **Expected results** (given independence of the two types of errors):

Wh-exhaustivity errors		0	1	2+
Universal errors	1	171 (84.9%)	23 (11.2%)	1 (3.9%)
	2	300 (84.9%)	40 (11.2%)	2 (3.9%)

(27) **Observed results**:

Wh-exhaustivity errors		0	1	2+
Universal errors	1	129	47	26
	2	211	80	62

As we can see, the number of exhaustivity errors made by children who made one or more

quantificational errors is much higher than expected (about twice the predicted rate for those who made one exhaustivity error and 25-30 times the predicted rate for those who made two or more exhaustivity errors. In order to test the correlation statistically, Strauss conducted a chi-square test on the observed versus predicted results. He found the discrepancy to be highly statistically significant: $p < .00000000001$. This shows quite conclusively that there is a relationship between exhaustivity errors and errors on universal quantification.

Strauss did not offer a detailed explanation for the correlation, but suggested that perhaps exhaustivity in answering wh-questions involves universal quantification over individuals for whom the answer is true. Hence, a theory maintaining that children do not possess correct semantics for universal quantification would predict that children are not able to provide exhaustive answers to wh-questions. A similar line of reasoning was argued for in Roeper et al. 2005, who proposed a developmental timeline in which the child progresses from the initial stage when wh-words possess a feature [+specificity], which leads children to interpret wh-words as 'singleton' (requiring a 'mention one' answer) to the exhaustivity stage (via a "plural" stage, when 'mention some' answers are permitted). The acquisition of exhaustivity under this view is triggered by and directly involves universal quantification over individuals, becoming good at which, as we saw from Strauss's work, correlates with becoming good with exhaustivity.

My proposal provides another explanation for the reported relationship between exhaustivity and universal quantification. The flexible account of exhaustivity (as Beck and Rullman) and my approach to "perfectionist errors" would handle the correlation quite well. The reason why the same children make quantification errors and wh-question exhaustivity errors is because the tests asking children to judge the truth-value of a universally quantified sentence involve choosing an appropriately exhaustive answer to a wh-question, a choice requiring pragmatic sophistication lacking in children this age.

A natural objection to raise at this point is that the quantificational errors discussed by Strauss included 'classic spreading' errors. He did not mention whether the corpus he

studied contained any perfectionist errors. This, however, does not mean that his conclusion about the relationship between q-spreading and exhaustivity is irrelevant for my theory of perfectionist errors. This is so because as I have discussed earlier children who give perfectionist responses also make errors in the 'extra object' condition. In fact, they make these errors at a much higher rate than the "classic spreading" children. In the next section I will discuss the puzzle of why the perfectionist children make classic spreading errors and why the rate of such errors is higher that the rate of classic spreading errors of the other group. I will suggest that what looks like the same type of error should be analyzed differently form the treatment I proposed for 'classic spreading' in chapter 3, i.e. the errors made in the 'extra object' condition by the group Philip called 'symmetry children'.

5. Why do "perfectionist" children make more "symmetrical" errors than the "symmetrical" children?

We have to explain why "perfectionist" children make significantly more errors in the 'extra object' condition as compared with the other group. I suggest that both errors are indicative of essentially the same phenomenon – a failure to identify correctly what is relevant. One way to account for this is to say that the classic spreading error for the perfectionist children is the same as for the other group – an error in specificity of the indefinite. The elevated rate of the error for the perfectionist group may be explained by this group's preference for singleton indefinites that carries over from *a* < *every* to the *every* < *a* sentences. This would suggest that the perfectionist group is the same type of children that rejected the 'object-universal' sentences in the inverse scope contexts in Experiment 1. This hypothesis needs further testing. However, there may be another explanation, which has to do not with children's interpretation of quantifiers, but with

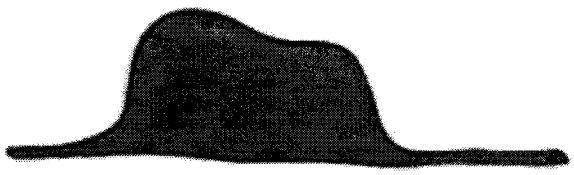
their interpretation of the task, as discussed above, and of what they see in the picture.

This explanation was suggested by anecdotal evidence from children's explanations for "symmetrical" responses in earlier research, as well as in my own studies. In a pilot study that I conducted, the child played the role of the puppet and had to describe asymmetrical pictures using universally quantified sentences. Her responses were quite illuminating. When shown a picture with four dogs each wearing a hat and an extra hat, she offered the following description: "Every dog is wearing a hat except one dog." When asked which dog was not wearing a hat, she pointed to the extra hat and explained "He didn't want to wear his hat because it looks silly." When asked again which dog was it, she said "he is gone".

This explanation points out an interesting property, directly related, in my view, to the Theory of Mind issues. While the adults look strictly at the surface of the picture, children construct a rich interpretation, including in it elements that are not overtly present in the picture, but clues for which they may see the picture contain. Thus, just as with "classic spreaders", these children read the salience of the single extra hat as a clue for its direct relevance for the task they are performing. They, however, differ in their strategy of how they treat the extra object. While classic spreaders choose to see the relevant single individual as a singleton member of the domain restrictions for the indefinite object, the perfectionist children construct a non-specific indefinite with multiple hats in its domain restriction. However, since the extra hat is seen by them as relevant, they have to assign a role it plays in the scenario they are evaluating. This role is seen by these children as standing for more than just a hat, but for a dog that failed to wear a hat.

This phenomenon was alluded to in the early literature on q-spreading as the Theory of a Missing Cowboy (Freeman, Sinha & Stedmon 1982) referring to the erroneous rejections of the sentence "every cowboy is riding a horse" because the picture contains an extra horse understood by the child as evidence of a missing cowboy. This reveals a TOM-related failure on the part of a child to notice that without overt linguistic context, an observer wouldn't assume any "behind-the-scenes" participants or events. This phenomenon was alluded to by de Saint Exupéry in the opening of the Little Prince, which contains an insight into the property of children's cognition that we now call their developing Theory of Mind and that may be responsible for children appearing 'illogical' to adults:

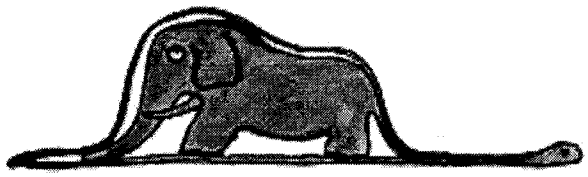
"I pondered deeply ... over the adventures of the jungle. And after some work with a colored pencil I succeeded in making my first drawing. My Drawing Number One. It looked something like this:



I showed my masterpiece to the grown-ups, and asked them whether the drawing frightened them.

But they answered: "Frighten? Why should any one be frightened by a hat?"

My drawing was not a picture of a hat. It was a picture of a boa constrictor digesting an elephant. But since the grown-ups were not able to understand it, I made another drawing: I drew the inside of a boa constrictor, so that the grown-ups could see it clearly. They always need to have things explained. My Drawing Number Two looked like this:



The grown-ups' response, this time, was to advise me to lay aside my drawings of boa constrictors, whether from the inside or the outside, and devote myself instead to geography, history, arithmetic, and grammar. That is why, at the age of six, I gave up what might have been a magnificent career as a painter. I had been disheartened by the failure of my Drawing Number One and my Drawing Number Two. Grown-ups never understand anything by themselves, and it is tiresome for children to be always and forever explaining things to them." Antoine de Saint Exupéry, The Little Prince

In this passage, de Saint Exupéry incisively points to the TOM-related phenomenon I believe is involved in 'perfectionist' children on the extra object condition - namely adults' failure to appreciate the child's predilection for 'seeing the unseen' and the child's Theory of Mind related failure to notice that an observer may not be expected to make the same assumption about objects not visible in the picture (plus a failure to realize that an irrelevant object is made highly perceptually salient for testing purposes, as a distracter).

To sum up, according to this hypothesis, the "perfectionist" responses are due to children's preference for an exhaustive response, which prompts them to reject any sentence that leaves out any participant that the child deems relevant. If the extra individual in an array of otherwise perfectly matched agent-theme pairs is of the same sort as the as the theme, children simply interpret it as representing a missing agent.

Thus, for a perfectionist child, a sentence 'every farmer is riding a donkey' uttered in the 'extra donkey' condition is not appropriate because it fails to mention the single donkey, which is seen by the child as relevant for the exhaustive description of the picture. The relevance of the single donkey, shown side by side with farmers riding donkeys, is seen as an indication that it is a stand-in for a farmer that is not riding it. This

would explain a nearly consistent rejection of these sentences by the perfectionist group in the "extra donkey" condition. For the 'classic spreading' group the error is a result of selecting which set to choose as relevant – the set of multiple farmer-donkey pairs or the single donkey (hence their oscillating between yes- and no- at the rate of 53-47%, for the perfectionist child everything in the picture is equally relevant and the single donkey is seen as part of the farmer/donkey relation (since no alternative reading here is considered, the rate of spreading responses is a high 82%, as reported in Philip 1995).

Chapter 6 What Does Theory of Mind Have to Do with This?

The project undertaken in this dissertation was aimed at addressing the important question of how generative theory should deal with the well-documented inconsistencies between child language and the properties of the adult grammar. One approach to this question attributes the pre-adult properties of child language to neurological maturation. A competing approach seeks to handle linguistic development within the Strong Continuity and Strong Uniformity assumption by accommodating any inconsistencies between the hypothesized grammar and the target grammar as incorrect or incomplete parameter setting without positing any ad hoc stages or arbitrary changes in linguistic development. Within this approach, any remaining discontinuity is handled by attributing it to performance factors, such as children's pragmatic or processing inflexibility, which may obscure the full extent of their adult-like grammatical competence.

The phenomenon of child language that I have focused on here, namely the so-called "q-spreading", has been argued to constitute evidence of discontinuity – representing a stage at which children possess certain grammatical mechanisms that are not found in any adult grammar, and which the child eventually 'grows out of'. The responses elicited from children may seem to indicate that children impose such truth conditions on the universally quantified sentences that require their semantic structure to include non-UG elements, namely restrictors formed by a conjunction of NPs, in which only one forms a syntactic constituent with the quantifier or even by the elements that are not mentioned in the sentence at all.

However, as shown by previous research, the seemingly odd truth conditions arise under predictable conditions and can be manipulated by the experimental technique. A

sharp decrease of errors can be produced by various types of manipulations, discussed in chapter 1. Yet, one generalization seems to hold across all experiments: children's performance with respect to q-spreading is facilitated by providing richer context (either linguistic or visual). As I have argued, the role of enriched context is that it supplies values for quantifier domain restrictions; in other words provides information as to what elements present in the context are relevant.[1] According to my proposal, the source of children's difficulty and the area in which development takes place is not their semantic competence, but the area in which meaning and context intersect. Since, as I discussed in chapter 3, the meaning of quantified sentences is partially determined by context, a child who has full semantic knowledge of universal quantification and scope ambiguities gets into trouble when she is asked to interpret such sentences with insufficient context. The specific conclusions I have arrived at are the following: I have claimed that children are adult-like in:

i) formal semantics of quantification;
ii. allowing all and only those scope relations between quantifiers that are permitted in adult grammar;
iii. knowing that quantifier domain restrictions are contextually determined;
iv. knowing that exhaustivity of wh-questions is (at least in part) context-dependent.

I have also claimed that children are different from adults in:

a) making mistakes in deciding what is relevant to the speaker in sub-optimal pragmatic conditions;
a) allowing the value of the covert domain restrictions to be set with subsets of the given individuals with a salient feature, even if the context is too poor for such domain narrowing for adults;
b) allowing singleton domains for the quantifiers that adults do not allow.
c) figuring out the intended Question under Discussion and using less than optimal context in deciding correct level of exhaustivity for the answer it requires.

[1] and in some cases clarifying that the speaker does not need to be exhaustive when describing the picture.

I have made a further claim, which requires clarification. I have suggested that the reason why children show the behaviors enumerated in the second list is their developing 'mind-reading' abilities, i.e. their ability to represent with some degree of accuracy the content of other minds in a given situation. More specifically, they lack in those skills that are necessary for negotiating the relationship between each interlocutor's (including self's) contribution to the discourse and information discernable by others in a given context.

That there may be a connection between pragmatic tasks in general and ToM is not surprising. Understood broadly, pragmatics deals with those areas of interpreting an utterance that involve aspects of meaning going beyond its truth conditions, namely various ways of using context in order to make a correct hypothesis about the speaker's meaning (i.e. reconstructing the proposition that the speaker intended to convey with both explicit and implicit information). Consider the following examples of this (from Wilson 2005):

1) Peter left *the party.* (political group or festive gathering)
2) The teachers told the students *they* (the teachers or the students) needed more time to finish the task.
3) I saw *no one* in town. (*no one I knew* or *no one interesting*)
4) *Some* of the talks were interesting (not all of them)

In order to decide what proposition the speaker intended to convey, the listener may have to lexically disambiguate (as in 1), establish co-reference (as in 2), fix quantifier domains (as in 3), or understand intended implications (as in 4) to list just a few examples of pragmatic tasks involved in sentence interpretation. Other examples of this involve recognizing context- based conditions on the felicitous use of an expression (presupposition accommodation) and discerning speaker's communicative intentions for

calculating conversational implicatures. All these tasks require the hearer to have a sophisticated ability to reason about the content of other minds. This connection between pragmatics and 'mind-reading' goes back to Grice (Grice 1957, 1967, 1969, 1989). Thus, his 'working out schema' for deriving conversational implicatures involves making a series of conjectures about the speaker's desire/belief psychology. Following this schema, in order for speaker A to infer the implicated meaning of B's utterance, A has to go through a very complex reasoning process involving a sophisticated ability to "read" other minds given in (6):

(5) A: *Is Sally coming to the meeting?*
B: *Her car broke down.* ⇒ Sally is not coming at the meeting.

(6) He said that P; he could not have done it unless he thought that P; he knows (and knows that I know that he knows) that I will realize that it is necessary to suppose that Q; he has done nothing to stop me from thinking that Q; so he intends me to think, or is at least willing for me to think, that Q. (Grice, 1989: 30-31)

There is a wealth of independent evidence that children undergo development with respect to their 'mind reading' abilities. It is fair to say that some questions within the field of the Theory of Mind development have not been settled, including what set of abilities precisely constitutes this cognitive module (or whether it is a module in the Fodorian sense or rather something that Fodor considered a central inferential cognitive system). However, it is widely accepted that there is a developmental schedule according to which various pieces of ToM come on line, including evidence that children's ability to make accurate inferences about epistemic effects of particular situations on others does not become adult-like until at least the age of 8. As I discussed in chapter 1, Taylor 1988 demonstrated that children until this age exaggerate how much information can be inferred about a mostly obscured picture by a naive observer. If this weakness

generalizes from visual information to communicative situations, it would predict that children may not be able to cope with the types of tasks illustrated above. Thus, working out conversational implicatures relies not just on the speaker's ability to know that the listener will be able to compute the implicated meaning in a given situation, but on the listener's ability to infer that the speaker knows that the listener would know that he knows that the implicated meaning can be computed.

Other aspects of meaning of sentences in context are derived by the listener from considering a set of alternative options the speakers had for expressing his intentions. For instance, scalar implicatures rely on the hearer considering possible alternatives, and assessing which alternative entails another and is thus more informative. Then, given the assumption that speakers would be maximally informative (or that the context is such that being maximally informative is required), the hearer would infer that the speaker has avoided the more informative option because his intention is to express an implicated meaning – negating the alternative that wasn't used.

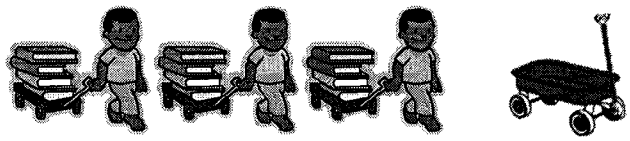
It is not surprising then that children have been shown to fail in computing conversational implicatures (e.g. Noveck 2001, Gedalyovich 2002). These results coexist with those findings that show that with improved experimental design, e.g. making the speaker's communicative intention to be maximally informative explicit, children are able to compute scalar implicatures (Papafragou and Tantalou 2005).

Children were also shown to use presuppostional expressions (e.g. definite determiners) in contexts in which their presupposition was not satisfied and couldn't be accommodated by the listener (Karmiloff-Smith 1988, Maratsos 1979). However, in addition to the infelicitous use of presuppositional expressions, there is evidence from

comprehension tasks that children's semantics of expressions such as definites is presuppositional (Maratsos 1979, Syrett, Kennedy, Lidz 2007) and that their problem in production tasks is not in knowing that the semantics of the expression requires a certain context in order to be interpreted, but in misjudging how much the listener can infer in a given situation in order to accommodate the presupposition. Thus, young children may use a definite NP in a context in which there is more than one object of the sort denoted by the NP. This presumably happens because they assume that what they perceive as unique (e.g. a red car in an array containing black, brown and white cars), would also be perceived as unique by the listener without realizing that the domain restriction has to be made explicit.

Now let's return to the cases of classic q-spreading and see whether we can make a similar more explicit connection between these responses and reasoning about the content of other minds. I have proposed that verifying sentences like (7) involves making judgments about how the domain of the indefinite should be restricted since the meaning of the sentence relies on whether the plural or the singleton indefinite is chosen.

(7) Puppet: Every boy is pulling a wagon.

(8) 

(9) Listener: Does he mean this? Or this one?

If the singleton option is chosen, then (7) receives the 'wide scope' indefinite interpretation and would be judged false in the context of (8). Presumably, both adults and children go through this type of reasoning, but adults quickly abandon the singleton

option for object indefinites unless there is overt context requiring them to choose it.[2] Many children, on the other hand, are almost as likely to go with this option as with the other one. I suggest that what causes the difference between adults and children is having to use a 'working out schema' that requires one to consider the content of other minds, something along the following lines:

(10) The speaker could mean the multiple X's or the single X for the domain restriction for the indefinite. Since the singleton X is more marked, if she means the single X, she has to assume that I know that she means the single X. However, she has no reason to assume that I would know that she means the single X since there have not been any overt clues given to me to inform me that she means the single X. Therefore, she must mean the multiple X's.

If we are correct in claiming that such reasoning is involved in making this decision, it is entirely plausible to suppose that q-spreading would disappear as children fully develop their ToM. Until that happens, they are unable to complete the inference in (10) and resort to guessing (or perhaps some other pragmatic strategy, such as using visual prominence of the objects in the picture as their guide).

Further research is necessary in order to demonstrate more directly the ToM/ Q-spreading connection. Perhaps, the Taylor-task (judging whether a partially obscured picture can be reconstructed by another observer) can be a good non-verbal predictor of a child's success with universal quantifiers.

If my conclusion about the ToM and children's pragmatic competence is correct, this means that we do not only have strong empirical reasons, reviewed in chapter 5, to adopt a pragmatic account of q-spreading that I have developed here, but we have a good theoretical justification for favoring such an account. It allows us to maintain the Strong Continuity Assumption, which provides a strong explanatory force to the acquisition

[2] Psycholinguistic evidence that adults consider both interpretations, but by default select the non-singleton option for object indefinites comes from Kurtzman and McDonald (1993).

theory. It also provides an explanation for 'why pragmatics should be different' and why, unlike semantics, it is subject to maturation.

Bibliography

Abusch, Dorit (1994). The scope of indefinites. *Natural Language Semantics* 2:83-136.

Beck, Sigrid and Hotze Rullman (1999). A fexible approach to exhaustivity in questions. *Natural Language Semantics* 7: 249-298.

Beghelli, Filippo, and Tim Stowell (1997). Distributivity and negation: The syntax of each and every. In *Ways of Scope Taking*,. Anna Szabolcsi (ed), 71–107. Dordrecht: Kluwer.

Bergsma-Klein, W. (1996). Specificity in Child Dutch: An Experimental study, Utrecht University.

Borer, Hagit and Kenneth Wexler (1992). Bi-uniqueness relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory*10, 147-189.

Bos, Petra, Bart Hollebrandse and Petra Sleeman (2004). Introduction: the pragmatics-syntax and semantics-syntax interface in acquisition; in *International Review of Applied Linguistics in Language Teaching* 42, 101-110; Walter de Gruyter.

Boster, Carole and Stephen Crain (1993). On children's understanding of *Every* and *Or. Proceedings of Early Cognition and Transition to Language*, University of Texas at Austin.

Brooks, P. J., and M. D. S. Braine, (1996). What do children know about the universal quantifiers *all* and *each? Cognition, 60*, 235-268.

Brooks, P. J., Jia, X., Braine, M. D., & da Graca Dias, M. (1998). A cross-linguistic study of children's comprehension of universal quantifiers: A comparison of Mandarin Chinese, Portuguese and English. *First Language, 18*(52), 33-79.

Bucci, W. (1978). The interpretation of universal affirmative propositions. *Cognition, 6*, 55-57.

Butterworth, G. 1991. The Ontogeny and Phylogeny of Joint Visual Attention; in *Natural Theories of Mind: Evolution, Development, and Simulation of Everyday Mindreading,* edited by Whiten, A.; Oxford: Basil Blackwell; p. 223-232

Chien, Y-C. and Kenneth Wexler (1987). Children's acquisition of the locality condition for reflexives and pronouns. Papaers and Reports on Child Language Acquisition, (PRCLA), Stanford university, California.

Chierchia, Gennaro. 2001. A puzzle about indefinites. In Semantic interfaces: Reference, anaphora and aspect, ed. Carlo Cecchetto, Gennaro Chierchia, and Maria Teresa Guasti, chap. 2, 51–89. Stanford, CA: Center for the Study of Language and Information.

Chierchia, Gennaro, Stephen Crain, Maria Teresa Guasti, and Rosalind Thornton (1998). "'Some' and 'or': A study on the emergence of Logical Form." *Proceedings of the 22nd Boston University Conference on Language Development*, 97-108.

Chomsky, Noam. 2001. Beyond Explanatory Adequacy. MIT Occasional Papers in Linguistics 20.

Crain, Stephen (2000). Sense and sense ability in child language. In *Proceedings of the 24th Boston University Conference on Language Development.* Somerville: Cascadilla Press.

Crain, Stephen, and Cecile McKee (1985). The acquisition of structural restrictions on anaphora. In *Proceedings of NELS 15*, 94-110. GLSA, University of Massachusetts, Amherst.

Crain, Stephen, Luisa Meroni and Andrea Gualmini (2000). The universal quantifier is strong in children, a paper presented at the 24th Annual Penn Linguistics Colloquium, University of Pennsylvania, Philadelphia, PA.

Crain, Stephen, and Rosalind Thornton (1998). *Investigations in Universal Grammar.* Cambridge, Mass: MIT Press.

Crain, Stephen, Rosalind Thornton, Carol Boster, Laura Conway, Diane Lillo-Martin, and Elaine Woodams (1996). "Quantification without qualification." *Language Acquisition* 5(2), 83-153.

Diesing, M. (1992). *Indefinites.* Cambridge, MA: MIT Press.

Drozd, Kenneth (2000) 'Children's weak interpretations of universally quantified questions'

Drozd, Kenneth (2001). Children's weak interpretations of universally quantified questions. In M.B. a. S. C. Levinson (Ed.), *Language Acquisition and Conceptual Development* (pp. 340-376). Cambridge: Cambridge University Press.

Drozd, Kenneth, and Erik van Loosbroek (1998). "Weak quantification, plausible dissent, and the development of children's pragmatic competence." In *Proceedings of the 23rd Boston University Conference on Language Development,* 184-195. Somerville: Cascadilla Press.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22:1-25.

Felber, Sarah (2002) Acquisition of two alleged polarity items: The convergence of children and adults in *some* and *any* judgments. Ms., University of Connecticut.

Filik, Ruth, Kevin Patterson and Simon P. Liversedge (2004). Processing doubly quantified sentences: Evidence from eye movements. *Psychonomic Bulletin and Review* 11 (5), 953-959.

von Fintel, Kai (1994). Restrictions on Quantifier Domains. PhD Dissertation, Graduate Student Linguistics Association (GLSA), University of Massachusetts, Amherst.

Fodor, Jery(1992). A theory of the child's theory of mind. *Cognition* 44, 283-296.

Fodor, Janet D. and Sag, Ivan (1982). Referential and quantificational indefinites. *Linguistics and Philosophy,* 5: 355-398.

Frazier, Lyn (1999) *On Sentence Interpretation.* Dordrecht: Kluwer Academic Publishers.

Freeman, N.H., C. G. Sinha, & J.A. Stedmon (1982) 'All the cars – which cars? From word meaning to discourse analysis'. In M.Beveridge (eds.), Children thinking through language. London: Edward Arnold, 52-74

Geurts, Bart. (2003). Quantifying kids. *Language **Acquisition**, 11*(4), 197-218

Geurts, Bart (2000). Indefinites and choice functions. *Linguistic Inquiry* **31**. 731-738.

Gordon, Peter (1996). "The Truth Value Judgment Task." In *Methods for Assessing Children's Syntax,* ed. Dana McDaniel, Cecile McKee and Helen Smith Cairns, 211-231. Cambridge, Mass: MIT Press.

Grice, Paul (1975). Logic and conversation. In Peter Cole and James Morgan (eds.) Syntax and Semantics 3: Speech Acts. New York: Academic Press, pp. 41-58.

Groenendijk, Jeroen and Martin Stokhof (1984). Studies on the Semantics of Questions and the Pragmatics of Answers. PhD Dissertation, University of Amsterdam.

Gualmini, Andrea (2003a). *The Ups and Downs of Child Language.* Doctoral dissertation, University of Maryland.

Gualmini, Andrea (2003b). Children do not lack *some* knowledge. *University of Maryland Working Papers in Linguistics* 12: 49-74.

Gualmini, Andrea and Stephen Crain (2001). "Downward Entailment in Child Language." In *University of Maryland Working Papers in Linguistics*, *11*, 112-133, University of Maryland at College Park.

Gualmini, Andrea, Stephen Crain, Luisa Meroni, Gennaro Chierchia and Maria Teresa Guasti (2001) At the Semantics/Pragmatics Interface in Child Language. *Proceedings of Semantics and Linguistic Theory 11*, 231-247. Ithaca, NY: Cornell University.

Gualmini, Andrea and Stephen Crain (2002). "Why No Child or Adult Must Learn De Morgan's Laws." *Proceedings of the 26th Boston University Conference on Language Development*, 342-254.

Gualmini A., L. Meroni and S. Crain (2003). An asymmetric universal in child language. In *Proceedings of the conference "Sub7– Sinn und Bedeutung"*, *Arbeitspapier* 114, Matthias Weisgerber (ed.). University of Konstanz, Germany.

Gualmini, Andrea, Luisa Meroni and Stephen Crain (2003). Children's asymmetrical responses, *Proceedings of 3rd Tokyo conference of Psycholingistics*. Hitsuji Shobo, Tokyo, Japan.

Hamblin, Charles Leonard (1973). Questions in Montague English. Foundations of Language 10: 41-53 [Reprinted in *Montague Grammar,* ed. Barbara H. Partee, 247-259. New York: Academic Press.].

Heim, Irene (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral dissertation, University of Massachusetts, Amherst.

Heim, Irene (1983). File change semantics and the familiarity theory of definiteness. In *Meaning, Use and the Interpretation of Language*, 164-190. Berlin: Walter de Gruyter. Reprinted in Portner and Partee, eds., 2002, 223-248. [Reprinted in Portner and Partee, eds., 2002, 223-248]

Heim, Irene (1983). On the projection problem for presuppositions. In *Proceedings of WCCFL 2*, ed. by D. Flickinger, 114–125. Stanford, Calif., CSLI

Hintikka, Jaako (1986). The Semantics of *A Certain. Linguistic Inquiry* 17:331-336.

Horn, Laurence (1989). *A natural history of negation.* Chicago, IL: University of Chicago Press.

de Hoop, Helen and Irene Krämer (2006). Children's optimal interpretations of indefinite subjects and objects, ms. Radboud University, Nijmegen.

Hurewitz, F., Brown-Schmidt, S., & Thorpe, K. J. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journalof Psycholinguistic Research, 29*(6), 597-626.

Hulsey, Sarah, Hacquard, Valentine, Fox, Danny and Gualmini, Andrea (2004). The Question-Answer Requirement and Scope Assignment. *MIT Working Papers in Linguistics Plato's Problems: Papers on Language Acquisition* 48, *71-* Massachusetts Institute of Technology

Hyams, Nina (1986). *Language acquisition and the theory of parameters.* Drdrecht: D. Reidel.

Inhelder, Barbel and Jean Piaget (1964). *The early growth of logic in the child.* London: Routledge, Kegan and Paul. Kramer, Irene (2000). Interpreting Indefinites. An

Experimental Study of Children's Language Comprehension. Doctoral dissertation, MPI Series. Wageningen: Ponsen & Looijen.
Kamp, Hans (1981). A theory of truth and semantic representation. In *Formal Methods in the Study of Language; Mathematical Centre Tracts 135*, eds. J.A.G. Groenendijk et al., 277-322. Amsterdam: Mathematical Centre. [Reprinted in: Jeroen Groenendijk, Theo Janssen, and Martin Stokhof (eds.), 1984, *Truth, Interpretation, Information*, GRASS 2, Dordrecht: Foris, pp. 1-41. Reprinted in Portner and Partee, eds., 2002, 189-222]
Karmiloff-Smith, Annette. 1979. *A Functional Approach to Child Language: A Study of Determiners and Reference*. New York: Cambridge University Press.
Karttunen, Lauri. 1977. 'Syntax and semantics of questions'. In *Linguistics and Philosophy* 1. 3-44.
King, Jeffrey (1988). Are Indefinite Descriptions Ambiguous? *Philosophical Studies* 53: 417-440
Krämer, Irene (2000). Interpreting Indefinites: an experimental study of children language comprehension. PhD dissertation. University of Utrecht.
Krämer, Irene (1999). Who Are You Talking about? Children's Interpretation of Sentence Initial Indefinite Subjects in Sentences Containing Negation. Linguistics in the Netherlands 16, 121-134.
Krämer, Irene (1998). Children's Interpretations of Indefinite Object Noun Phrases: Evidence from the Scope of Negation. Linguistics in the Netherlands, 15, 163-174.
Kratzer, Angelika. 1998. Scope or pseudo-scope? Are there wide-scope indefinites? In *Events in Grammar*, ed. S. Rothstein, 163-196. Dordrecht: Kluwer.
Kratzer, Angelika. 2004. Covert quantifier domain restrictions. A talk at Milan Meeting. Palazzo Feltrinelli, Gargnano.
Kurtzman and MacDonald, 1993. Resolution of quantifier scope ambiguities. *Cognition*, (48): 243 – 279.
Ladusaw, William (1979). *Negative Polarity Items as Inherent Scope Relations*. Unpublished Ph.D. Dissertation, University of Texas at Austin.
Lee, T. H. T. (1986) Studies on Quantification in Chinese. Doctoral Dissertation, UCLA.
Lidz, Jeffery and Julien Musolino (2002). Children's command of quantification. *Cognition* 84: 113-154.
Lillo-Martin, Diane (1991). *Universal Grammar and American Sign Language: setting the null argument parameters*. Dordrecht: Kluwer Academic Press.
Maratsos, Michael (1976). The Uses of Definite and Indefinite Reference in Young Children: An Experimental Study of Semantic Acquisition. Cambridge: Cambridge University Press.
Maratsos, Michael, Dana Fox, Judith Becker and Mary Ann Chalkley (1985). Semantic restrictions on children's passives. Cognition 19: 167-191.
Marti, Luisa (2003). Contextual Variables. Doctoral dissertation, University of Connecticut.
Matthewson, Lisa (1999). On the Interpretation of Wide-Scope Indefinites. *Natural Language Semantics* 7:79-134.
May,Robert (1977). The Grammar of Quantification. Doctoral dissertation, Massachusetts Institute of Technology.

Meroni, Luisa. 2005. *Putting Children in Context.* University of Maryland, College Park Dissertation.

Meroni, Luisa, Andrea Gualmini and Stephen Crain (2000). "A conservative approach to quantification in child language." *Proceedings of the 24th Annual Penn Linguistics Colloquium*, 171-182.

Meroni, Luisa, Stephen Crain and Andrea Gualmini (2001). "Felicity Conditions and On-line Interpretation of Sentence with Quantified NPs, paper presented at the *14th Annual CUNY Conference on Human Sentence Processing*, March 15th-17th, 2001. Philadelphia, University of Pennsylvania.

Miller, K. and C. Schmitt. 2003: Wide-scope Indefinites in English Child Language. Utrecht

Miller, Karen and Cristina Schmitt (2003). The interpretation of Indefinites and bare singulars in Spanish child language. Paper presented at the 6th conference on the acquisition of Spanish and Portuguese as first and second languages. University of New Mexico, Albuquerque.

Milsark, Gary (1977). Toward an explanation of certain peculiarities of the existential construction in English. Linguistic Analysis 3(1): 1-29.

Musolino, Julien (1998). Universal Grammar and the Acquisition of Semantic Knowledge. Doctoral dissertation, University of Maryland.

Musolino, Julien, Stephen Crain and Rosalind Thornton (2000). Navigating negative quantificational space. Linguistics 38(1): 1-32.

Musolino, Julien and Jeffrey Lidz (in press). Why children aren't universally successful with quantification. To appear in Linguistics.

Musolino, Julien and Jeffery Lidz (2003). The scope of isomorphism: turning adults into children. Language Acquisition 11(4): 272-291.

Noveck, I. (2001). When children are more logical than adults: Experimental investigations of scalar implicatures. *Cognition, 78*, 165-188.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition,86*, 253-282.

Papafragou, Anna and Niki Tantalou (2004). Children's computation of implicatures. Language Acquisition 12(1), 71-82.

Philip, William (1995). *Event quantification in the acquisition of universal quantification.* Doctoral dissertation, University of Massachusetts, Amherst.

Philip, William (1996). The Event Quantificational Account of Symmetrical Interpretation and a Denial of implausible Infelicity." In *Proceedings of the 20th Boston University Conference on Language Development,* 564-575.

Pinker, Stephen (1984). *Language Learnability and Language Development.* Cambridge, MA: Harvard University Press.

Radford, Andrew (1990). *Syntactic theory and the acquisition of English syntax.* Oxford: Basil Blackwell.

Rakhlin, Natalia. To appear. Children's Knowledge of the Negative Determiners. Generative Approaches to Language Acquisition in North America 2, McGill University, Canada.

Rakhlin, N. 2004: A New Approach to Quantifier-Spreading. Proceedings of the 23rd West Coast Conference on Formal Linguistics. V. Chand, A. Kelleher, A. J. Rodríguez, and B. Schmeiser (eds.): 635-648

Reinhart, Tanya (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20: 335-397.

Roeper, Tom and Ed Matthei (1975) 'On the **acquisition** of all and some.' *Papers and Reports on Child Language Development 9.* Stanford: Stanford University Press

Roeper, Tom and Jill de Villiers (1991). The Emergence of Bound Variable Structures. In T. L. Maxfield and B. Plunkett, (eds.), *Papers in the acquisition of wh.* GLSA, University of Massachusetts, Amherst.

Roeper, Tom, Uri Strauss and Barbara Zurer Pearson (2004). The Acquisition Path of Quantifiers: Two Kinds of Spreading, UMass Amherst, on-line ms.

Russell, Bertrand. 1948. *Human knowledge: Its scope and limits.* London: Allen

Scaife, M. and J.S. Bruner. (1975), The Capacity for Joint Visual Attention in the Human Infant, *Nature* 253: 265-266

Schwarz, Bernard. 2001. Two kinds of long-distance indefinites. In van Rooy and Stokhof (2001), 192–197.

Schwarzschild, Roger (2002). Singleton Indefinites, *Journal of Semantics,* 19: 289-314.

Spelke, E. S., & Cortelyou, A. (1981). Perceptual aspects of social knowing: Looking and listening in infancy. In M. E. Lamb & L. R. Sherrod (Eds.), Infant social cognition: Empirical and theoretical considerations. Hillsdale, NJ: Erlbaum.

Su, Yi-Ching (2001). Scope and specificity in child language: a cross-linguistic study of English and Chinese. PhD Dissertation. University of Maryland, College Park.

Strauss, Uri. 2006 to appear. 'The Acquisition of Exhaustivity'. In Tanja Heizmann (ed) *UMOP (University of Massachusetts Occasional Papers) 33.* University of Massachusetts, Amherst.

Sugisaki, Koji and Miwa Isobe (2001). Quantification without Qualification without Plausible Dissent, SULA, UMass

Szabolsci, A. (1997). *Ways of **Scope** Taking.* Dordrecht: Kluwer.

Tanenhaus, Michael, Michael Spivey-Knowlton, Kathleen Eberhard and Julie Sedivy (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* 286: 1632-1634.

Taylor, Marjorie (1988). The development of children's understanding of the seeing knowing distinction. Janet W. Astington, Paul L. Harris, David R. Olson (eds.) *Developing Theories of Mind.* Cambridge University Press: 207-225. Cambridge, England.

Trueswell, John C., Irina Sekerina, Nicole M. Hill and Marian L. Logrip (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition* 73: 89-113.

de Villiers, Jill and Helen Tager Flusberg (1975). Some facts one simply cannot deny. *Journal of Child Language* 2: 279-286.

Wason, Peter (1972). In real life negatives are false. Logique et Analyse 15: 17-38.

Wilson, Deirdre (2005). New directions for research on pragmatics and modularity. *Lingua* 115: 1129-1146.

Wimmer, Heinz, Juergen Hogrefe, and Beate Sodian (1988). A second stage in children's conception of mental life: Understanding informational accesses as origins of knowledge and belief. Janet W. Astington, Paul L. Harris, David R. Olson (eds.)

Developing Theories of Mind. Cambridge University Press: 173-192. Cambridge, England.

Winter, Yoad (1997). Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20: 399-467.

Westerstahl. Dag (1985). Logical Constants in Quantifier Languages, *Linguistics and Philosophy* (8), 387-413.

Yatsushiro, K. and U. Sauerland. 2005: The Acquisition of Presuppositions: The Presuppositions of Every. Paper presented at the conference "Experimental Pragmatics: Exploring the Cognitive Basis of Conversation". Centre for Research in the Arts, Social Sciences and Humanities, University of Cambridge.