Quantifying kids

Bart Geurts

University of Nijmegen

Abstract

It has been known for several decades that young children have difficulties with universal sentences. In this paper I present an analysis of the main errors that have been reported in the literature. My proposal is based on an old idea, viz. that children's errors are caused by a non-canonical mapping from syntactic form to semantic representation. Previous theories based on this assumption were not entirely successful because they lacked the proper framework for dealing with quantification. In particular, they failed to recognise the importance of the distinction between weak and strong quantifiers. Everything falls into place, or so I argue, if this distinction is taken into account.

1. The facts

Stories about language acquisition often begin in Geneva, and this one is no exception. Inhelder and Piaget (1959) presented children with displays of coloured squares and circles, and asked their subjects to assess quantified sentences against these displays. They found that the task is a lot harder than one should suspect. Here are a few examples:

(1) Scene: 14 blue circles, 2 blue squares, 3 red squares.

    Q: Are all the circles blue?
    A: No, there are two blue squares.

(2) Scene: blue circles, blue squares, red squares [no exact numbers given].

    Q: And are all these squares red?
    A: No, because there is a blue one.
    Q: And are all the blue ones circles?
    A: Yes.

(3) Scene: 5 blue circles, 3 red squares.

    Q: Well, look, are all the circles here blue?
    A: Yes ... no.
    Q: Why?
    A: There are red ones.
    Q: Where?
    A: There are red squares and blue circles.

Similar results have been obtained in countless subsequent experiments, which demonstrated beyond any reasonable doubt that young children often deviate from their elders when it comes to evaluating quantified sentences. Although most experimental work has been done with English-speaking subjects, comparable data have been reported for Dutch (Philip and Verrips 1994, Philip and Coopmans 1995, Drozd and van Loosbroek 1999), Turkish (Freeman and Stedmon 1986), Japanese (Takahashi 1991), Catalan (Philip 1995), Chinese (Chien and Wexler 1989), and of course French (Inhelder and Piaget 1959). Furthermore, Philip and Avrutin (1998) present data showing that agrammatic aphasics may suffer from the same problems.

    With the help of an artificial example, the non-adult responses found in the literature may be classified as follows. When asked if the sentence 'Every X is Y' is true or false, a child may claim that the sentence is:

(A) false if $\|X\| = \{a,\ b,\ c\}$ and $\|Y\| = \{a,\ b,\ c,\ d\}$

(B) true if $\|X\| = \{a,\ b,\ c,\ d\}$ and $\|Y\| = \{a,\ b,\ c\}$

(C) false if $\|X\| = \{a,\ b,\ c\}$, $\|Y\| = \{a,\ b,\ c\}$, and $\|Z\| = \{d\}$

(Where $\|X\|$ is the extension of X in a given context, i.e. the set of individuals of which X holds.) According to this classification (1) counts as a Type-A response, which puts it in the same class with (4) and (5):

(4) Scene: 5 apples and 3 pigs eating 1 apple each.

    Q: Every pig is eating an apple ... Does this picture go with the story?

    A: No. Those two apples have no pig.   (Philip and Takahashi 1991)

(5) Scene: 4 garages and 3 cars, each occupying 1 garage.

    Q: All the cars are in the garages.

    A: No.   (Donaldson and Lloyd 1974)

Inhelder and Piaget's example (2) betrays a Type-B error, as does the following:

(6) Scene: 5 cars and 4 garages, each occupied by 1 of the cars.

    Q: All the cars are in the garages.

    A: Yes.   (Donaldson and Lloyd 1974)

Type-C errors are harder to find (in the literature, that is), but example (3) is a case in point, as is the following:

(7) Scene: 3 cats holding a balloon, and 1 mouse holding an umbrella.

    Q: Is every cat holding a balloon?

    A: No. (pointing to the mouse)   (Philip and Verrips 1994)

Problems with quantified sentences have been observed in 3-year olds and persist at least up to age 7. Not all children have them, but many of them do, though it is hard to say how many, because reported error rates vary greatly; this variation makes it pointless to give precise figures, but error rates in excess of 50% are quite common. It is unlikely that all types of error are equally frequent, but unfortunately the literature doesn't yield a clear distribution pattern, because experimenters have tended to confine their attention to Type-A and Type-B errors (and the former category seems to

have been especially popular). Indeed, Type-C responses have been used to screen out subjects in pretests, so there is little in the way of systematic data on this category. Nonetheless, such evidence as is available suggests that Type-C errors are much rarer and less persistent than others.

Children's problems with quantified sentences involve a number of factors. One important factor is the type of quantifier. To begin with, errors seem to be restricted to sentences with universal quantifiers: Takahashi (1991) reports that sentences with cardinal quantifiers don't cause any trouble, and the same goes for sentences with definite subjects (Drozd 2001), unless they contain 'floated' universal quantifiers (Donaldson and Lloyd 1974). Smith (1979, 1980) found a clear contrast between 'some' and 'all', sentences with 'some' nearly always prompting adult-like responses, and according to Braine and Rumain (1983), the only problem with 'some' is that it occasionally has numerical connotations, leading young children to reject 'Some X are Y' on the ground that the number of X's is 'too large', but such findings are irrelevant in the present context. Hence, the problems described above are confined to sentences with 'all', 'every', and 'each'. There may be differences amongst these quantifiers, as well, but as on this score the empirical record is neither very substantial nor entirely consistent, they will be ignored in the following.[1]

Not only is there a clear distinction between tasks with universal and existential sentences: problems arising from the former may be compounded by previous exposure to the latter. Smith (1979, 1980) presented 4- to 7-year-olds with quantified questions like 'Are all animals cats?', which had to be resolved not against a visual display but against basic world knowledge (of a kind available already to the youngest subjects). Half of the children started with the batch of 'all' questions, while the other half was first taken through the 'some' questions. Smith's main results were that the first group performed quite well on all tasks, while the second group had

---

[1] Drawing on unpublished research by himself, Krämer, and Loosbroek, Drozd (2001) reports that Dutch five-year-olds recognise the semantic difference between 'alle' ('all') and 'iedere' ('every'), while four-year-olds don't. On the other hand, according to Drozd and Philip (1993) and Philip (1995), English-speaking children are indifferent to the distinction between 'all' and 'every'. On yet another hand, Freeman and Stedmon (1986) found differences between 'all', 'absolutely all', 'every', and 'every single', which only emerge with sufficiently large situations: with a 4-garage/3-car array the choice of quantifier doesn't have an effect, but with a 4-garage/5-car array it does. Clearly, this is an issue that calls for more experimental research.

considerable difficulties with 'all' questions (questions with 'some' proved to be unproblematic). It appears, therefore, that initial exposure to a series of 'some' questions may cause errors with subsequent 'all' questions, but not vice versa.

Another parameter affecting children's performance is the way a situation is laid out and presented. Compare, for instance, Donaldson and Lloyd's examples (5) and (6). It seems as if in these tasks the cars and garages are not on an equal footing: being more natural landmark objects, the garages appear to function as the background against which the cars are viewed. This initial impression is reinforced by Freeman et al.'s (1982) finding that a group's relative salience may bias children's responses one way or the other, a finding that was confirmed in experiments by Drozd and van Loosbroek (1999):

(8) Scene: 4 cowsheds and 3 cows, occupying 1 cowshed each; cowsheds have been made salient in the preceding discourse.
   Q: Are all the cows in the cowsheds?
   A: No.   (Freeman et al. 1982)

(9) Scene: as in (8), but now the cows are more salient.
   Q: Are all the cows in the cowsheds?
   A: Yes.   (Freeman et al. 1982)

In the same vein, Freeman and Stedmon (1986) discovered that group size may be a factor, too: further exploring the car/garage paradigm first introduced by Donaldson and Lloyd, Freeman and Stedmon found that 'the more cars, the more correct answers' (p. 41). A similar improvement may be achieved by keeping the collection not of garages but of cars constant across tasks (Freeman et al. 1982). Taken together, these data suggest rather strongly that, as compared to adults, children establish the domain of quantification with more regard to pragmatic clues and proportionally less regard to grammatical constraints. If a collection of individuals is particularly salient (for whatever reason), children tend to assume that a given quantifier ranges over it, despite the fact that grammatical constraints bar such an interpretation. This is not yet a full-fledged explanation, though, as it leaves open a number of questions. In particular, it remains to be seen what it means for a grammatical constraint on quantification to be less rigid in children than it is in adults, and why this holds for universal quantifiers only.

One way of describing the errors children make is that they tend to interpret a universal statement as if it entailed its converse; as if 'All the cows are in the cowsheds' entailed 'All the cowsheds are occupied by cows'. This is reminiscent of a well-attested error made by adults in deductive reasoning, which has come to be known as 'illicit conversion'. When Newstead and Griggs (1983) asked their adult subjects whether 'All Y are X' follows from 'All X are Y', about one third of them claimed that it did, and it is widely agreed that illicit conversion of universal propositions accounts for a fair share of the errors adults make in syllogistic reasoning (Newstead 1989, Geurts 2003). So adults have their problems with universal quantifiers, just as children have. But the problems are not the same. Adults make mistakes with *arguments* that involve universal quantification; they have no trouble whatsoever with the evaluation tasks used in acquisition studies. This is not to say that the two phenomena are completely unrelated, for I do believe that the ultimate cause is the same. It is that universal quantification is more complex than most other varieties. But this complexity does not seem to affect children and adults in the same way.

2. Full grammatical competence?

The research surveyed in the foregoing has been criticised by Crain and his associates for being based on a flawed experimental design (Crain et al. 1996, Crain and Thornton 1998). Crain et al. maintain that earlier studies failed to observe certain pragmatic felicity conditions, which led some children to respond in a way that is out of step with their linguistic competence. In particular, it is crucial, according to Crain et al., that in a truth value judgment task both answers should be made available as genuine options:

> In the contexts for yes/no questions, felicitous usage dictates that both the assertion and the negation of a target sentence should be under consideration. (p. 302)

As applied to example (4), say, this principle calls for a scenario in which the possibility that some pigs don't eat an apple is considered, too. If this possibility is left out of account altogether (as it usually would be in experiments prior to Crain et al.'s), children will be puzzled as to why the experimenter wants to know if the statement is true. Crain et al. report that in experiments satisfying this condition children performed about as well as adults,

and they conclude that 'young children have full grammatical competence with universal quantification.' (Crain et al. 1996: 83)

Crain and his co-workers are staunch nativists, who 'take it to be the null hypothesis that children have full linguistic competence' (Crain et al. 1996: 147). Someone taking a more neutral view on language acquisition will find their argument uncompelling, and will probably draw a more guarded conclusion, viz. that Crain et al. have hit upon yet another pragmatic influence on children's response patterns. But even this is disputable (cf. also Geurts 2000, Drozd 2001). To begin with, though it is beyond doubt that there are all sorts of felicity conditions which constrain linguistic processing, it remains to be seen that the one identified by Crain et al. is among them. Contrary to what Crain et al. contend, it is doubtful that a yes/no question is pragmatically infelicitous unless both the affirmative and the negative answer are 'under consideration' in any substantial sense. On the contrary, it is part of the function of yes/no questions to *introduce* alternatives into the discourse; whether or not they are already under consideration is immaterial.

In my own experience, children are rather good at answering all manner of questions that would be infelicitous according to Crain et al., and there is plenty of experimental evidence to confirm this impression. For example, as we saw in the foregoing, children have no difficulties assessing universally quantified sentences against basic world knowledge, nor are they troubled by sentences with definite or indefinite subjects. Not only do such data argue against Crain et al.'s felicity condition, they also remain unexplained on the theory they propose.

A further reason for doubting Crain et al.'s diagnosis is that the parameter they claim to have manipulated in their experiments is open to alternative construals. To see why, consider how (4) would have to be presented in order to be pragmatically felicitous, according Crain et al. This would call for a story which had one or more pigs consider whether or not they should eat an apple before all pigs finally decided to have one. Such a story would inevitably raise the salience of the three pigs, of course, and we knew already that salience may affect children's responses. Hence, it isn't even clear that Crain et al.'s main finding is a new one.

Finally, even if it were true that Crain et al.'s experiments prove children to have 'full grammatical competence' with respect to universal quantification, their diagnosis still leaves open some of the most intriguing issues raised by the empirical facts. Why is it, for example, that older children and adults aren't bothered by an experimental set-up that, by Crain et al.'s lights, is

7

hopelessly flawed? Crain et al. have very little to say about this:

> We believe the reason is that older children and adults are better at taking tests than young children. To be successful in previous studies, participants were required to accommodate the fact that the negation of the test sentences was not under consideration in the adult interpretation. Presumably, older children and adults have learned to see through misleading circumstances in which test sentences are presented [. . .]. (Crain et al. 1996: 117)

This is not very specific, and doesn't even begin to explain the data reviewed in Section 1. Or, to put the same point in a positive way: Crain et al.'s nativist views on budding linguistic competence leave room for a broad range of theories of how young children *process* universal sentences, including some proposals that have been made in the literature as well as the proposal to be advanced below. So, although I happen to disagree with Crain et al., their views are not necessarily inconsistent with the analysis to be presented below.

3. Quantification in discourse

In a nutshell, the theory I propose is this. Children's problems with universal quantification are due to a malfunctioning mapping from syntactic structure to semantic representation. If this mapping 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. I will argue that these pragmatic processes are the same in children and in adults, so even if their effects are different, the only thing that goes wrong, strictly speaking, is the syntax-semantics mapping.

In the semantic literature on quantification it is generally assumed that there is a basic distinction between strong and weak quantifiers. In English, the litmus test for separating the two classes is provided by existential 'there' sentences:

(10)   There are {three / at least five / no / a few / *all / *most} wombats in the sauna.

While the universal quantifiers and 'most' are strong, 'three', 'at least five', 'no', and 'a few' are weak. More accurately, the latter are weak by default, because they too are inadmissible in the following frame:

(11) *There are {three / at least five / no / a few} *of the* wombats in the sauna.

The key difference between strong and weak quantifiers is that the former but not the latter are inherently relational. 'Most X are Y' always means that most of the individuals in a given set of X's are Y; 'At least five X are Y' admits of an analogous interpretation, but does not require it, for it may also be construed as expressing merely that at least five individuals have the property of being X as well as Y. Following established usage, I will apply the 'strong/weak' terminology to quantifiers as well as their interpretations. Hence, a quantifier is strong if it always has a strong reading; weak quantifiers have a weak interpretation only by default.

The weak/strong distinction plays a central role in my proposal, which is why I will discuss it at some length. The discussion will be couched in the framework of discourse representation theory (Kamp 1981, Kamp and Reyle 1993, Geurts 1999), but DRT is not essential to the account I propose, which could be embedded in other frameworks as well, provided they can differentiate between strong and weak quantifiers along the lines to be set out in the following.

DRT consists of two main components: a language of (mental) semantic representations (SRs) and a so-called 'construction algorithm', which maps incoming sentences (or rather, their syntactic analyses) onto SRs. The SR-language is defined by a simple syntax, which is given a truth-conditional interpretation. In the following I will explain the intended meanings of SRs in informal terms; a more technical description is given in the appendix.[2]

I am going to assume that the semantic representations of weak (construals of) quantifiers are different from strong ones. Here is an example with a weak quantifier:

(12) a. Fred photographed two llamas.
     b. ⟨two⟩[x: llama(x), Fred photographed x]

The intended interpretation of (12b) is that there are two individuals x such

---

[2] A caveat for readers familiar with DRT: for the purposes of this paper I use only a fragment of the standard semantic representations, and will exhibit only the parts I need.

that x is a llama and Fred photographed x. More generally, if a quantifier Q is weak, it prompts the introduction of an SR of the form $\langle Q \rangle \varphi$. Speaking somewhat loosely, the meaning of such a represention is that there are Q-many individuals that satisfy φ. The semantics of weak quantifiers given in the appendix entails that, in an SR of the form $\langle Q \rangle \varphi$, Q binds the first variable in φ; any remaining variables have existential force by default.[3]

If a quantifier is strong, it demands a relational interpretation. For example, (13a) claims that all individuals in a given set of llamas—the domain of the quantifier—have a certain property, viz. they were photographed by Fred. In order to represent this reading, we need the SR in (13b):

(13) a. Fred photographed all llamas.
      b. [x: llama(y)]$\langle$all$\rangle$[Fred photographed x]

The intended interpretation of an SR of the form $\varphi\langle$all$\rangle\psi$ is that all individuals of which φ is true satisfy ψ, as well. φ represents the quantifier's 'domain', i.e. the set of individuals the quantifier ranges over; ψ represents the 'nuclear scope', or 'scope' for short. The semantics of strong quantifiers given in the appendix entails that, in an SR of the form $\varphi\langle Q \rangle\psi$, Q binds the first variable in φ; any remaining variables have existential force by default.

Example (13a) illustrates how the domain of a (nominal) quantifier is constrained by the grammar: the quantified NP 'all llamas' perforce ranges over llamas. The grammar of English dictates that this be so, but it is not the only way quantifier domains are constrained, for there are pragmatic factors, as well. To begin, a sentence like (13a) would normally be used in a context in which some particular set of llamas is already given. In general, a strongly quantified NP 'Q X' is like an anaphoric pronoun in that it prompts a search for a contextually salient collection of X's, which are to serve as Q's domain. Or, to put it the other way round, the context constrains the domain of a quantifier by making salient this or that set of individuals.

There is another way in which pragmatic inferences narrow down quantifier domains, which is exemplified by (14a):

(14) a. Most people visit Berlin [in the spring]$_F$.

---

[3] This regime of variable binding deviates somewhat from the standard DRT convention. I introduce a different method here because I will be claiming that the grammatical connection between a quantifier and its domain of quantification is less rigid in children than it is in adults, and this idea is more difficult to implement with the standard treatment of variable binding.

b. Most people who visit Berlin (at some point in time), do so in the spring.

Part of the VP in (14a) is focused, while the remainder, i.e. 'visit Berlin', is backgrounded (non-focused) material. Since both focused and backgrounded material are part of the VP we should expect both to go into the quantifier's nuclear scope. However, as the paraphrase in (14b) suggests, while the focused content behaves as expected, the backgrounded material actually ends up restricting the domain of the quantifier. That is to say, the interpretation we would expect on purely grammatical considerations is (15a), in which t ranges over time intervals, and backgrounded material is underlined; the observed interpretation is (15b), in which the backgrounded material has moved to the domain of the quantifier:

(15) a. [x: person(x)]⟨most⟩[t: <u>x visits Berlin during t</u>, spring(t)]
      b. [x, t: person(x), x visits Berlin during t]⟨most⟩[spring(t)]

I have argued elsewhere that this remarkable phenomenon is best explained in terms of presupposition (Geurts 1999, Geurts and van der Sandt 1999). For the purposes of the present discussion, that is as it may be. The important thing for now is just *that* the focus/background structure of a sentence may restrict the domain of a quantifier, too.

We have seen that there are two types of pragmatic constraints on quantifier domains. On the one hand, quantifiers prefer to have domains that are contextually salient; on the other hand, the division between focus and background plays a part in the process of domain restriction, too. There is a two-way interaction between these factors, in the sense that the interpretative clues provided by salience and focusing should mesh, as they do in (16a) but fail to in (16b) and (16c).

(16) Berlin always attracts lots of visitors, but . . .
      a. most people visit Berlin [in the spring]$_F$.
      b. ?most people [visit Berlin]$_F$ in the spring.
      c. ?most [people]$_F$ visit Berlin in the spring.

If a collection of individuals X is salient, and X is assigned to be Q's domain, then one should expect X to be backgrounded. And vice versa, if X is backgrounded, and the domain of Q is a salient part of the discourse, then that part would normally consist of X's.

4. Weak construals of strong quantifiers

To return to the main topic, my proposal is as follows. Children who give non-adult responses to universal sentences construe the (strong) universal quantifier as if it were weak: the problem lies in the mapping between syntactic form and semantic representation; it is a parsing problem. More accurately, it starts out as a parsing problem, which is repaired by pragmatic processes of the kind discussed in the last section.

I assume that the distinction between weak and strong quantifiers is not just a linguistic curiosity, but is directly relevant to the ways quantified expressions are processed. This assumption is at the heart of the analysis of quantification outlined above, according to which strong quantifiers must be represented by relational structures, while weak quantifiers give rise to non-relational representations by default; a weak quantifier calls for a relational representation only if it triggers a domain presupposition, as in the following example:

(17)   Some (of the) llamas resented being photographed by Fred.

If this is understood as referring only to the llamas owned by Fred's wife (say), 'some llamas' gets a strong construal, and therefore has a relational representation. However, this is a marked case; as a rule, 'some' will be weak.[4]

It follows from this treatment of weak and strong construals that the former are simpler than the latter. Strong readings are harder to obtain not only because their semantic representations are more intricate, but also because they require a more roundabout mapping from form to meaning (cf. Figure 1). When processing a sentence of the form 'Q X are Y', X and Y must be processed separately if Q is strong; whereas if Q is weak, X and Y may be interpreted as if they were conjoined.

From a semantical point of view, the crucial factor is that weak quantifiers are intersective while strong quantifiers are not: in order to ascertain whether 'Q X are Y' is true, if Q is weak, we only need to inspect the set $\|X\| \cap \|Y\|$. Hence, the set of lawyers that are crooks is all we need for checking if 'Some lawyers are crooks' is true; any crooks or lawyers outside $\|lawyer\| \cap \|crook\|$ need not be taken into account. Of course, this does not hold anymore if we

---

[4] (17) is marked in at least two respects: it has an indefinite subject (subjects are almost always definite) and requires an accent in an unusual position, i.e. on 'some'.
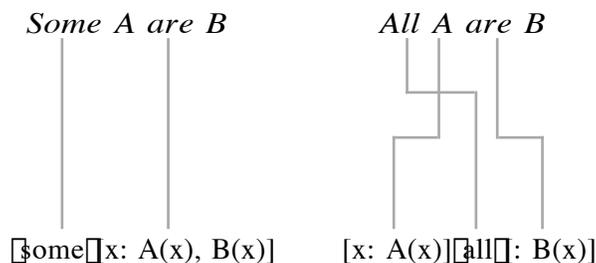
Figure 1: Mapping surface form to semantic representation: 'some' vs. 'all'

replace 'some' with a strong quantifier such as 'all', for strong quantifiers are non-intersective. Therefore, it is not a coincidence that existential sentences afford simpler representations than universal ones: at root, the difference is a matter of content.

The proposed differentiation between existential and universal quantifiers is supported not only by linguistic data but by experimental findings, as well. It has been known for some time that existential sentences are easier to process than universal ones. For example, in an experiment by Just (1974), sentences of the form 'Some X are Y' had shorter response latencies than sentences of the form 'All X are Y', except when $\|X\|$ and $\|Y\|$ were identical or disjoint, in which cases universal and existential sentences were processed equally fast; the distribution of errors followed the same pattern, i.e. subjects made more mistakes with universal sentences. The same regularities had been observed previously by Meyer (1970) in a rather different experimental set-up, which goes to show that they are fairly robust.

The foregoing considerations show that weak quantifiers can be handled with simpler means than strong ones, and that it is plausible to assume that they are. But I cannot claim to have proved that it is so, because it is always possible in principle to generalise to the worst case, and uniformly treat weak and strong quantifiers as giving rise to relational interpretations. That is to say, in DRT terms, it is possible to map all quantified sentences onto semantic representations of the form $\varphi\langle Q\rangle\psi$, regardless whether Q is strong or weak. However, this strategy will not work unless we emend our model of interpretation in other ways as well. In particular, we will have to drop the assumption that a relational quantifier presupposes its domain by definition, and stipulate instead that domain presuppositions are triggered on strong construals only. Thus, strong construals turn out to be more

13

complex than weak ones, after all, although the difference doesn't register in terms of grammar and/or parsing. As far as I can see, there is no knock-down evidence against this line of analysis. However, it seems to me that an account such as mine is more plausible, precisely because it assumes that in general simpler representations and simpler processing strategies will be preferred. Generalising to the worst case allows for a model that economises on the *diversity* in representations and processing strategies, which is a way of saving, too, though it is arguably less urgent from a processing point of view.

The theory I propose starts out from the assumption that the child's grasp of the grammar of universal quantification is not deficient in any way. Nor do I see any compelling reason for doubting that children master the logic of universal quantification. In my view, the child's problem (if there is one) lies in the mapping from grammatical form to semantic representation, which is more complicated for universal quantifiers than it is for others. Children have a certain tendency to interpret all quantifiers as if they were weak, because it is easier to do so. Thus, if the weak processing strategy is applied to (18a), the sentence is mapped onto (18b) instead of the adult representation in (18c):

(18) a. Every boy is riding an elephant. (Drozd and van Loosbroek 1999)
     b. ⟨every⟩[x, y: boy(x), elephant(y), x rides y]
     c. [x: boy(x)]⟨every⟩[y: elephant(y), x rides y]

(18b) hardly counts as a full-blown construal of (18a), of course, because 'every' is a relational quantifier, and I assume that the child knows this; that is to say, the quantifier's lexical meaning is transparent enough, it is just the mapping from form to meaning that goes awry. Hence, the child's target representation is something like (19):[5]

(19) [... : ...]⟨every⟩[x, y: boy(x), elephant(y), x rides y]

(19) is the child's semantic representation before pragmatic processing sets in. This representation leaves the domain of quantification underdetermined (as compared to the adult's construal), so that there is more elbow-room for pragmatic inferencing.

In Section 3 we saw how focus and salience conspire to restrict quantifier

---

[5] The transition from (18b) to (19) may seem ad hoc, but it isn't, because an analogous transformation is needed for weak quantifiers to get strong construals.

domains. Suppose now that this is the same in children as in adults. Then the domain of (19) will be constrained by the focus/background division within the nuclear scope, which in its turn is constrained by the context. If the child considers that a given set of boys is currently the most salient discourse entity, he will assume it is the intended domain and that [x: boy(x)] is backgrounded, and will therefore interpret (18a) as (18c). Notwithstanding the fact that it was obtained by a non-canonical procedure, this interpretation is correct by adult standards. If, on the other hand, the child's attention is focused on a given set of elephants, [y: elephant(y)] is backgrounded, and the same sentence is interpreted as (20):

(20)   [y: elephant(y)]⟨every⟩[x: boy(x), x rides y]

The reading in (20) will prompt a Type-A response if every boy rides an elephant and some elephants are *sans* boy; it will prompt a Type-B response if every elephant is ridden by a boy and some boys are *sans* elephant; and it will produce the correct response if it so happens that all elephants are ridden by a boy and all boys are riding an elephant.

It will be clear how this analysis accounts for the facts discussed in Section 1. First, it explains why children's errors are restricted to sentences with universal quantifiers: this is so because universal quantifiers require a more intricate mapping from form to meaning; weak quantifiers are easier to process. Secondly, it explains why errors with universal sentences may be caused by previous exposure to existential sentences: this will happen if the processing strategy for weak quantifiers, which is easier anyway, is primed by a series of existential sentences, and carried over to subsequent universal sentences. Thirdly, it explains how and why pragmatic reasoning contributes to children's construals of universal sentences: pragmatic reasoning can play a larger role in children than in adults whenever an incorrect mapping from surface form to semantic representation leaves the quantifier's domain underdetermined; but the pragmatic mechanisms that take over at this point are the same for all ages.

Thus far I haven't said anything about Type-C errors. How are they accounted for? This is a hard question, though not for lack of possible answers; it is just that the dearth of data doesn't really allow us to distinguish good answers from bad ones. For this reason, the following remarks are quite tentative. One way of looking at Type-C responses is that they are no different from Type-A and Type-B responses. To explain how, let us go back

to example (7), which I repeat here for convenience:

(21)   Scene: 3 cats holding a balloon, and 1 mouse holding an umbrella.
       Q: Is every cat holding a balloon?
       A: No. (pointing to the mouse)

Suppose that a child presented with this scene homes in not on the animals and objects it contains but rather on the fact that everyone is holding something. Applying the same reasoning as in the foregoing, such a child (if it failed to get the right mapping from grammatical form to semantic representation) would arrive at the following interpretation:

(22)   [x, y: x holds y]⟨all⟩[cat(x), balloon(y)]

This says that every individual that holds something is a cat holding a balloon, which accounts for the Type-C error in (21).

Appealing though this line of analysis may be, I don't think it is right, for two reasons. First, and most importantly, it is at odds with the well-established fact that children are strongly object-oriented in the sense that they tend to concentrate their attention on medium-sized physical objects. Only to give one illustration of this bias, in an experiment conducted by Shipley and Shepperson (1990), 3- to 6-year-olds were presented with toy ducks in various colours, and asked to count the number of different colours, in response to which many children, but especially the younger ones, counted the ducks instead of the colours (see Bloom 2000 for further discussion and references). Results like this are hard to reconcile with the assumption that a child who gives a Type-C response, as in (21), finds the relation of holding more salient than the individuals in the scene. An additional reason for dismissing the current proposal is that it doesn't really explain why Type-C errors are so much rarer and less persistent than others.

I have the impression that the error in (21) is due to the fact that the child quantifies over all animate individuals in the scene (compare the use of the indefinite pronoun 'ones' in (3), which apparently refers to all figures on display, although the target sentence quantifies over circles only). That is to say, the child in (21) seems to obtain something like the following reading:

(23)   [x: animal(x)]⟨all⟩[y: cat(x), balloon(y), x holds y]

If this is right, the problem is that the descriptive material furnished by

the sentence doesn't constrain the domain of quantification in any way, be it grammatically or pragmatically. It should be expected, therefore, that children who make Type-C errors have the same mapping problem as children who make Type-A or Type-B errors, and in addition have an insecure grasp of the pragmatic principles of interpretation discussed in Section 3. This line of explanation seems promising to me, because it does justice to the intuition that Type-C errors are clearly different from, and more serious than, the others, but I must stress once more that in the absence of more systematic data, devising hypotheses tends to be a somewhat gratuitous exercise.

5. Comparisons with other accounts

The chief merit I would claim for my proposal is that it gets the facts right without having to appeal to ad hoc assumptions. But what is nice about it, too, I think, is that it links up to several ideas that have been around in the literature for some time, and allows us to say that there is something right about many alternative accounts. In the present section I will comment on some of these connections.

My analysis locates the problem with universal quantification in one particular area, viz. the syntax/semantics mapping. In this respect, I agree with Donaldson and Lloyd (1974) and Bucci (1978), among others, who have pointed in the same direction. According to Bucci, for example, problems with universal sentences arise because semantic relations between words fail to be properly encoded:

> The [universal] sentence is encoded as a simple string or un-ordered set of substantive words without hierarchical structure [. . .]. This 'structure-neutral' form is initially registered simply as a listing of semantic information including the main content words, e.g., 'all, blue, circles'. Then some further interpretation may or may not be imposed, determined by context or by guessing strategy—not by the original sentence structure as such. (Bucci 1978: 58-59)

It will be clear that my proposal is very much in the same spirit. But there are differences, too. For one thing, my account is more explicit than what Bucci proposes. For another, I don't share Bucci's assumption that the semantic

17

representations implicated in children's non-adult responses are 'structure-neutral'. In my view, there is really nothing amiss with these representations as such or with the processes involved; it is just that they are not the appropriate ones for universal sentences. Hence, the process of interpretation is considerably more constrained than Bucci suggests, and this holds not only for its initial phase, but all the way. For, on my account, it is definitely wrong to say that children arrive at an interpretation by mere guesswork: the pragmatic mechanisms that produce Type-A and Type-B responses are quite standard, and only for Type-C responses may it be necessary to relax the assumption that children have adequate mastery of the relevant pragmatic principles.

Another line of approach that bears some resemblance to mine has it that children are liable to misconstrue universal quantifiers as adverbs of quantification. This is an interesting idea because, in general, adverbials impose relatively weak constraints on their domain of quantification, which makes them more sensitive to pragmatic influences, hence more susceptible to misconstruals of the type I have been discussing. Adverbial theories have been proposed by Roeper and de Villiers (1993) and Philip (1995), among others; I will briefly discuss the latter here. According to Philip, children tend to construe universal quantifiers as quantifying over events rather than individuals:[6]

> [. . .] children, for whatever reason, tend to assign to a sentence containing a single determiner universal quantifier [sic] the logical form which adults assign to a sentence containing an adverb of quantification, for which the default mode of quantification is quantification over events. (Philip 1995: 53)

Transposing Philip's analysis into the DRT framework (and simplifying it somewhat), the idea is that (24a) is parsed as (24b):

(24) a. Every boy is riding an elephant. (= (18a))
      b. [e: . . .]⟨every⟩[x, y: boy(x), elephant(y), x is riding y in e]

This construal leaves the domain of quantification underdetermined, and calls for further constraints. These are furnished, in Philip's proposal, by a set

---

[6] Philip doesn't explain why this should be the case, although he tentatively suggests that quantification over events may be more basic than quantification over individuals. This is unlikely, however, in view of children's general object bias (cf. Section 4).

of special-purpose rules, which might yield something like the following, for example:

(25)   [e, y: elephant(y), y participates in e]⟨all⟩[x: boy(x), x is riding y in e]

(The details of this representation deviate considerably from the official version, but it is only the principal features of Philip's analysis that I am interested in here.) Broadly speaking, this is similar to what I have proposed: the two accounts agree that a hitch in the grammatical analysis produces a quantifying structure whose domain is virtually unconstrained. The main differences are, first, that according to Philip the child resorts to an entirely different style of quantification, which I find implausible, and secondly, that the mechanisms which, in Philip's analysis, determine the domain of quantification are plainly ad hoc.

Although Drozd's (2001) slogan is the same as mine ('Universal quantifiers are construed as if they were weak'), the resemblance between our accounts doesn't go much further than that. Drozd explicitly rejects the notion that the locus of the problem lies in the mapping from form to meaning, thereby suggesting that the child's misunderstanding is rather deeper. But if the syntax/semantics mapping is not the problem, what is? Strangely enough, Drozd doesn't say: he doesn't define what it means for a universal quantifier to be interpreted as a weak one, which makes his proposal somewhat elusive. (It is rather as if someone were to claim that '+' may be construed as denoting a unary operation without saying which one.) Fortunately, however, Drozd explains at some length what is supposed to follow from his proposal, and it is on his corollaries that I will focus in the following.

The logical inferences licensed by weak quantifiers deviate systematically from the inference patterns associated with strong quantifier. For example, if Q is weak, 'Q X are Y' is equivalent to 'Q X that are Y, are Y'. Thus, 'Some lawyers are crooks' is true iff 'Some lawyers who are crooks, are crooks' is true.[7] Drozd maintains that this equivalence is the source of Type-B errors. If a child interprets 'every' as weak and hence intersective, he may infer that (24a) is true iff (26) is, and proceed to verify the latter instead of the former, thus arriving at the conclusion that (24a) is true in a situation in which not all boys ride an elephant.

---

[7] It is true that the latter has an air of redundancy about it, but we are interested here in truth-conditional content only.

(26)   Every boy who is riding an elephant is riding an elephant.

This proposal raises a number of questions that I shall not go into (though it should be noted that (26) is tautologous, and therefore immune to falsification). Let us turn instead to Drozd's views on Type-A errors. Given the way Drozd handles Type-B errors, one should expect this class to be treated along the same lines, for if 'all' is interpreted as a weak quantifier, it must be symmetric as well as intersective, and (24a) must be equivalent not only to (26) but also to:

(27)   Every elephant is being ridden by a boy.

A child latching on to this equivalence would commit a Type-A error, evidently.[8]   However, Drozd ignores this possibility, and argues that Type-A responses arise in a different way altogether.

What causes Type-A errors, according to Drozd, is the fact that weak quantifiers are context-sensitive in distinctive ways, and it is this context sensitivity that allows for Type-A construals of universal quantifiers. The key piece in Drozd's argument is Westerståhl's (1985) example:

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

Westerståhl observes that, on its most likely interpretation, (28) means that the number of Scandinavian Nobel prize winners is larger than one would expect, statistically speaking. Drozd's idea is that this somehow carries over to universal sentences, as interpreted by Type-A children:

> As in Westerståhl's account of the preferred in interpretation of [(28)], children who make this error may arrive at a response to the question *Is every boy riding an elephant?* by first comparing the number of elephant-riding boys with what they consider to be the normal or expected frequency of elephant-riders given the situation shown in the picture. If children use the presence of the extra elephant [...] to infer that the expected or normal frequency is four [while there are only three boys riding an elephant], they will say *no* [...] (Drozd 2001: 358-359)

---

[8] Furthermore, Type-C errors should be in for a similar treatment, since it holds for weak quantifiers that 'Q X are Y' is equivalent to 'Q D are X and Y', where D names the universe of discourse.

Note, however, that the parallel Drozd discerns between the standard construal of (28) and the Type-A construal of (24a) is imperfect, at best, because it is hard to see how expected frequencies can interact with the interpretation of 'every' in anything remotely like the way they do with the interpretation of 'many'. Note, furthermore, that 'many' and 'few' are in fact unique even among the weak quantifiers in their sensitivity to expected frequencies. So strictly speaking Drozd's thesis is not that Type-A errors are caused by weak construals of universal quantifiers, but rather by a mistaken assimilation of 'every' and its kin to the 'many' type. Why do children fall into this curious mistake? Drozd doesn't say. Without going further into the details of Drozd's proposal, this should suffice to show that, despite our verbatim agreement about what is the main problem, his analysis is rather different from mine.

6. Learnability

A general problem with the theories discussed in the last section is that they have difficulties explaining the fact that, sooner or latter, virtually all children become adept at universal quantification. For example, according to adverbial theories such as Philip's, a child has to negotiate an entirely different style of quantification in order to master the intricacies of nominal quantifiers like 'every' and 'all', and quite apart from the question why this detour has to be taken in the first place, it is not so easy to see how the child achieves it. In a word, learnability is a problem for adverbial theories, and the same issue also arises, though in different ways, for the other theories discussed in the foregoing.

One of the advantages of the analysis presented in Section 4 is that it makes it relatively easy to see how children manage to overcome their problems with universal quantification. In fact, my proposal is compatible with many different stories about the acquisition of quantification, and as I suspect that it depends on one's ideological predispositions how much one appreciates stories like this, I will briefly relate two: one for nativists and one for empiricists.

The nativist story goes as follows. Let us suppose that everything the child needs to *know* for dealing with quantification is part of its biological heritage. We are supposing, therefore, that the syntax and semantics of quantification are innate. (I take it that this is the position taken by Crain

and his associates, more or less.) In terms of the processing model adopted here, this is to say that the child enters the world with fully developed systems of syntactic and semantic representation, as well as a complete battery of rules mediating between the two. If my proposal is correct, some of these mapping rules will be more complex than others. In particular, the rule that is supposed to deal with strong quantifiers is more involved than the weak-quantifier rule. Moreover, the semantic representation of a strong quantifier is more complex than that of a weak one. Therefore, limitations of working memory or attention will tend to affect the interpretation of strong quantifiers more than they affect the interpretation of weak quantifiers. Furthermore, since the triggering conditions for these mapping rules are similar (strong and weak quantifiers aren't that different, syntactically speaking), it is to be expected that, when things go wrong, the weak-quantifier rule will cut in, especially if we may assume that weak quantifiers are more frequent than strong ones, and therefore the mapping for the former is not only easier but also primed more strongly.[9]

If this story is on the right track, improvements in working memory and/or attention will automatically solve the child's problems with universal quantification, and there can hardly be any doubt that such changes do take place in early childhood. Basically, according to this story, what a child has to do in order to master universal quantification is grow up; which is what one should expect from a nativist account.

The prerequisites for my empiricist story are fewer, because it drops the assumption that mapping rules are innate; they will have to be acquired somehow.[10] So, part of the job of learning a language is to find a mapping

---

[9] As things currently stand in the semantic literature, quantifying expressions fall into three classes: 'definitely strong', 'definitely weak', and 'as yet undecided'. There is a consensus that the universal quantifiers and 'most' are strong. There is also a consensus that the following are weak: 'no', 'some', 'a few', 'at least n', 'more than n', 'at most n', 'less than n'. Finally, there are quantifiers like 'many' and 'few' whose status is moot, although their distribution pattern is that of the weak quantifiers (they may occur in existential 'there' sentences, for example). These observations suggest that the strong quantifiers are a minority, and support the conjecture that strong quantifiers occur less frequently than weak ones.

[10] Hard-nosed empiricists would say, of course, that syntactic and semantic representations must be acquired, too. I will not go as far as that, though mainly for expository reasons: if syntax and semantics can't be treated as given, either, my story becomes somewhat messy. (But then empiricist accounts generally *are* messier than their nativist competitors, precisely because they seek to unburden the child's genetic endowment.)

between syntax and semantics. My central claim is that in certain respects this mapping is simpler than in others. In particular, the mapping is simpler for weak quantifiers than it is for strong ones, and therefore one should expect the former to develop earlier, and, for the same sort of reasons that figure in the nativist story, to occasionally interfere with the workings of the latter.

The empiricist story presupposes that parsing and interpretation are, to some degree at least, independent systems. This runs counter to what is perhaps still the prevailing view in linguistics and psychology, namely that meaning depends entirely on syntax, and therefore syntactic processing must precede interpretation. If this were the case, it would be impossible to *find* a syntax-semantics mapping, simply because any semantic representation presupposes such a mapping. However, there is a lot of evidence that interpretation does not parasitise syntax in a very strict sense. There are experimental results suggesting rather strongly that semantics and syntax are interleaved processes (e.g. Marslen-Wilson et al. 1988, MacDonald et al. 1994, van Berkum et al. 1999), but even without evidence from the laboratory it is quite obvious that interpretation depends on contextual clues, world knowledge, and prior expectations, just as much as it depends on syntax.

Concluding remarks

The theory presented in this paper is predicated on the assumption that children's errors with universal quantification are due to a deficient syntax/semantics mapping. What is new about my approach is that it explains why this mapping should cause trouble, to begin with, and that it provides an explicit account of the comprehension processes underlying children's responses, which also shows how and why pragmatic factors interfere with the interpretation of universal sentences.

Even if this account is correct as far as it goes, it is certainly not the end of the story, if only because we badly need more data. For at present too little is known about Type-C errors, the influence of the collective/distributive distinction ('all' vs. 'every' and 'each'), and the longitudinal dimension of error patterns, to name only some of the topics that, in my opinion, deserve to be explored further.

Acknowledgment

References

Bloom, P. 2000: *How Children Learn the Meanings of Words.* MIT Press, Cambridge, Mass.

Braine, M.D.S. and B. Rumain 1983: Logical reasoning. In: J.H. Flavell and E.M. Markman (eds.), *Handbook of Child Psychology, Volume 3: Cognitive Development.* Wiley and Sons, New York. Pp. 263-340.

Bucci, W. 1978: The interpretation of universal affirmative propositions. *Cognition* 6: 55-77.

Chien, Y. and K. Wexler 1989: children's knowledge of relative scope in Chinese. *Stanford Child Language Research Forum.*

Crain, S., R. Thornton, C. Boster, L. Conway, D. Lillo-Martin, and E. Woodams 1996: Quantification without qualification. *Language Acquisition* 5: 83-153.

Crain, S. and R. Thornton 1998: *Investigations in Universal Grammar.* MIT Press. Cambridge, Mass.

Donaldson, M. and P. Lloyd 1974: Sentences and situations: children's judgments of match and mismatch. In: F. Bresson (ed.), *Problèmes Actuels en Psycholinguistique.* Presses Universitaires de France, Paris.

Drozd, K.F. 2001: Children's weak interpretations of universally quantified questions. In: M. Bowerman and S.C. Levinson (eds.), *Language Acquisition and Conceptual Development.* Cambridge University Press. Pp. 340-376.

Drozd, K.F. and W. Philip 1993: Event quantification in preschoolers' comprehension of negation. In: E.V. Clark (ed.), *Proceeding of the 24th Annual Child Language Research Forum.* Stanford. Pp. 72-86.

Drozd, K.F. and E. van Loosbroek 1999: Weak quantification, plausible dissent, and the development of children's pragmatic knowledge. *Proceedings of the 23rd Annual Boston University Conference on Language Development*, 184-195.

Freeman, N.H., C.G. Sinha, and J.A. Stedmon 1982: All the cars—which cars? From word meaning to discourse analysis. In: M. Beveridge (ed.), *Children Thinking Through Language.* Edward Arnold, London. Pp.52-74.

Freeman, N.H. and J.A. Stedmon 1986: How children deal with natural language quantification. In: I. Kurcz, G.W. Shugar, and J.H. Danks (eds.), *Knowledge and Language.* Elsevier, Amsterdam. Pp. 21-48.

Geurts, B. 1999: *Presuppositions and Pronouns.* Elsevier, Oxford.

Geurts, B. 2000: Review of Crain and Thornton (1998). *Linguistics and Philosophy* 23: 523-532.

Geurts, B. 2003: Reasoning with quantifiers. *Cognition* 86: 223-251.

Geurts, B. and R. van der Sandt, 1999: Domain restriction. In: P. Bosch and R. van der Sandt (eds.), *Focus: Linguistic, Cognitive, and Computational Perspectives.* Cambridge University Press. Pp. 268-292.

Inhelder, B. and J. Piaget 1959: *La Genèse des Structures Logiques Élémentaires: Classifications et Sériations.* Delachaux et Niestlé, Neuchâtel. English translation (1964): *The Early Growth of Logic in the Child: Classification and Seriation.* Routledge and Kegan Paul, London.

Just, M.A. 1974: Comprehending quantified sentences: the relation between sentence-picture and semantic memory verification. *Cognitive Psychology* 6: 216-236.

Kamp, H. 1981: A theory of truth and semantic representation. In: J.A.G. Groenendijk, T.M.V. Janssen, and M.B.J. Stokhof (eds.), *Formal Methods in the Study of Language.* Mathematical Centre Tracts 135, Amsterdam. Pp. 277-322.

Kamp, H. and U. Reyle 1993: *From Discourse to Logic.* Kluwer, Dordrecht.

MacDonald, M.C., N.J. Pearlmutter, and M.S. Seidenberg 1994: Lexical nature of syntactic ambiguity resolution. *Psychological Review* 101: 676-703.

Marslen-Wilson, W., C.M. Brown, and L.K. Tyler 1988: Lexical representations and language comprehension. *Language and Cognitive Processes* 3: 1-16.

Meyer, D.E. 1970: On the representation and retrieval of stored semantic information. *Cognitive Psychology* 1: 242-300.

Newstead, S.E. 1989: Interpretational errors in syllogistic reasoning. *Journal of Memory and Language* 28: 78-91.

Newstead, S.E. and R.A. Griggs 1983: Drawing inferences from quantified statements: a study of the square of oppositions. *Journal of Verbal Learning and Verbal Behavior* 22: 535-546.

Philip, W. 1992: Distributivity and logical form in the emergence of universal quantification. In: C. Barker and D. Dowty (eds.), *Proceedings of SALT II.* Ohio State University. Pp. 327-345.

Philip, W. 1995: *Event Quantification in the Acquisition of Universal Quantification.* Ph.D. thesis, University of Massachusetts, Amherst.

Philip, W. and S. Avrutin 1998: Quantification in agrammatic aphasia. In: U. Sauerland and O. Percus (eds.), *The Interpretive Tract.* MIT Press, Cambridge, Mass. Pp. 63-72.

Philip, W. and P. Coopmans 1995: Symmetrical interpretation and scope ambi-

guity in the acquisition of universal quantification in Dutch and English. Ms., University of Utrecht.

Philip, W. and M. Takahashi 1991: Quantifier spreading in the acquisition of *every*. T.L. Maxfield and B. Plunkett (eds.), *Papers in the Acquisition of WH*. GLSA, Amherst, Mass. Pp. 267-282.

Philip, W. and M. Verrips 1994: Dutch preschoolers' *elke*. Paper presented at the 1994 Boston University Conference on Language Development.

Shipley, E.F. and B. Shepperson 1990: Countable entities: developmental changes. *Cognition* 34: 109-136.

Smith, C.L. 1979: Children's understanding of natural language hierarchies. *Journal of Experimental Child Psychology* 27: 437-458.

Smith, C.L. 1980: Quantifiers and question answering in young children. *Journal of Experimental Child Psychology* 30: 191-205.

Takahashi, M. 1991: Children's interpretation of sentences containing *every*. In: T.L. Maxfield and B. Plunkett (eds.), *Papers in the Acquisition of WH*. GLSA, Amherst, Mass. Pp. 303-323.

Van Berkum, J.J., C.M. Brown, and P. Hagoort 1999: Early referential context effects in sentence processing: evidence from event-related brain potentials. *Journal of Memory and Language* 41: 147-182.

Westerståhl, D. 1985: Logical constants in quantifier languages. *Linguistics and Philosophy* 8: 387-413.

Appendix: Interpreting semantic representations

The semantic-representation language employed in this paper is a dialect of DRT (see Kamp and Reyle 1993 or Geurts 1999 for introductions). The SRs used to represent quantified sentences are of the form $\varphi\langle Q\rangle\psi$ and $\langle Q\rangle\varphi$, where φ and ψ consist of two components: a sequence of variables and a set of conditions. A DRT-style semantics interprets these structures with respect to a model $\mathbf{M}$ and a partial assignment function f. We say that f *embeds* an SR $[u_1 \ldots u_m: \gamma_1 \ldots \gamma_n]$ *into* $\mathbf{M}$ iff there is a function $g \supseteq f$ such that $\mathrm{dom}(g) = \mathrm{dom}(f) \cup \{u_1 \ldots u_m\}$, and g embeds all $\gamma_1 \ldots \gamma_n$ into $\mathbf{M}$.

To illustrate the interpretation of quantified SRs, I give the definitions for the weak quantifier 'at least five' and the strong quantifier 'all'. Let φ be of the form $[u_1 \ldots u_m: \gamma_1 \ldots \gamma_n]$; then:

(1) f embeds $\langle\text{at-least-five}\rangle\varphi$ into $\mathbf{M}$ iff there are at least five individuals a for which there is a $g \supseteq f$ such that $g(u_1) = a$ and g embeds φ into $\mathbf{M}$.

(2) f embeds $\varphi\langle\text{all}\rangle\psi$ into $\mathbf{M}$ iff for all individuals a for which there is a $g \supseteq f$ such that $g(u_1) = a$ and g embeds φ into $\mathbf{M}$, there is an $h \supseteq g$ that embeds ψ into $\mathbf{M}$.

Note that in both cases the quantifier binds the first variable in φ (i.e. u_1); any remaining variables undergo an existential interpretation.