

## Open Access

Hubert Haider\*, Christina Schörghofer-Essl and  
Karin Seethaler

## Quantifying kids prefer intersecting sets – a pilot study

DOI 10.1515/zfs-2017-0003

**Abstract:** Children between approximately four and five years of age are known to fail in picture matching tasks with verbal stimuli presenting an existentially quantified object NP in the scope of a universally quantified subject NP. In this paper, we suggest an experimentally tested provisional answer to a question that has not been asked in any previous work on the very phenomenon: Would they also fail for the *truth-conditionally equivalent* stimuli in which the universal quantifier is replaced by a negated existential quantifier (plus a negated predicate, as in *Every boy walks with a balloon* vs. *No boy walks without a balloon*).

The experimental results indicate that the latter task is indeed much easier. This lends support to the hypothesis that the primary source of the difficulties lies in the acquisition of a special aspect of compositional semantics, that is, in computing the semantic effects of a universal vs. an existential quantifier as subject in combination with an existential quantifier in the predicate. For a universally quantified subject, the relation between the restrictor set of the quantified subject and the set denoted by the predicate is one of set inclusion. For an existentially quantified subject, it is set intersection. At least for the picture matching tasks, the intersection relation seems to be easier to handle than the inclusion relation, even for semantically equivalent stimuli.

**Keywords:** acquisition of semantics, quantifier spreading, exhaustive pairing

---

**\*Corresponding author: Hubert Haider**, Fachbereich Linguistik / Centre for Cognitive Neuroscience, Universität Salzburg, E-mail: [hubert.haider@sbg.ac.at](mailto:hubert.haider@sbg.ac.at)

**Christina Schörghofer-Essl**, Fachbereich Linguistik, Universität Salzburg,  
E-mail: [christina.schoerghofer-essl@stud.abg.ac.at](mailto:christina.schoerghofer-essl@stud.abg.ac.at)

**Karin Seethaler**, Fachbereich Linguistik, Universität Salzburg,  
E-mail: [karin.seethaler@stud.sbg.ac.at](mailto:karin.seethaler@stud.sbg.ac.at)



© 2017 Hubert Haider et al., published by De Gruyter.

This work is licensed under the Creative Commons Attribution 4.0 License.

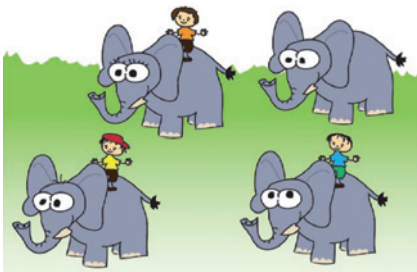
Angemeldet | [hubert.haider@sbg.ac.at](mailto:hubert.haider@sbg.ac.at) Autorexemplar

Heruntergeladen am | 05.06.17 16:56

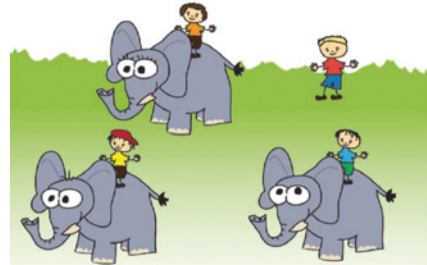
# 1 Introduction

The phenomenon has been detected by Inhelder and Piaget (1958). A generally accepted explanation is still wanting. The puzzle is this. In picture selection tasks with stimuli that involve universally quantified subjects combined with an indefinite (i.e. existentially quantified) object, the reactions of children of about four to six years of age differ from older children, whose reactions are adult-like.<sup>1</sup> In particular, when asked e.g. “*Is every/each boy riding an elephant?*” they tend to reject a picture such as in Figure 1.

When confronted with such a picture and inquired “*Every boy rides an elephant. Correct?*” typically well more than half of the children will respond “No” and, when asked why, they typically tell the experimenter that there is an additional elephant without a boy.



**Figure 1:** Example of a stimulus for situation (1).



**Figure 2:** Example of a stimulus for situation (2).

Figure 2 serves as a test case for adult-like “no”-reactions. When confronted with this picture and the same question, nevertheless a high percentage of children will wrongly respond “Yes”, as if the question (to a child that is able to correctly master this task) had been “*Every elephant carries/is ridden by/a boy. Correct?*” The surmise (cf. Drozd 2001) that children insist on exhaustive pairing of elephants and boys does not completely match with the fact that the percentage of mistakes triggered by this kind of stimulus picture is typically lower than in the Figure 1 setting.

The very same children would have no problems with existentially quantified utterances like “Some boys ride an elephant” or “Some elephants carry a

<sup>1</sup> Let us bear in mind in this connection that Brooks and Sekerina (2006: 199) report a mean percentage of only 79 % to 80 % correct responses for the adults they tested with Figure-1-type stimuli.

boy”.<sup>2</sup> Seethaler (2012) tested<sup>3</sup> *alle* ‘all’ vs. *jeder* ‘each’, both in adjacent and distant positions<sup>4</sup> and did not find any significant differences, between the two variants of quantifier items. The results were virtually identical.

In general, the test stimuli used in the literature have the form (1a). The kids then have to decide whether a given statement of the form (1a) is correct relative to the depicted situations (1b), as in Figure 1, or to a depicted situation (1c) as illustrated by Figure 2. The verbal stimuli and the respective pictures for (1b) and (1c) are presented to them at different times and randomized.

- (1) a. For every A-element x there is a B-element y, such that x is in a specified relation to y.
- b. Situation 1: several A-element items, all in relation to B-element items; one unrelated B-kind item (see Figure 1)
- c. Situation 2: each B-element item in relation to an A-element; one A-element item unrelated to a(ny) B-element (see Figure 2)

## 2 Competing accounts in the literature

The explanations for the failure patterns proposed in the literature take one of three alternative circumstances to be the major causal factor, namely (i) *semantic* deficiencies in composing the meaning of universally quantified sentences, (ii) *syntactic* deficiencies in handling NP structures with universal quantifiers, or (iii), *pragmatic* deficiencies (pragmatic felicity conditions of the test items, distraction by the test situation, with kids being bothered by the experimental set-up).

The main objection on pragmatic grounds is that the test situation presents utterances in a context that does not offer a chance of *plausible dissent* or *plausible denial* (Freeman and Stedmon 1986). “In the contexts for yes/no questions, felicitous usage is said to dictate that both the assertion and the negation of a target sentence should be under consideration.” (Crain et al. 1996: 102) If

<sup>2</sup> Smith (1979, 1980) confirms a clear contrast between *some* and *all*. Stimuli with *some* nearly always prompt adult-like responses.

<sup>3</sup> Twenty-one children; mean age 5;06 (youngest participant: 4;10; oldest one 7;04)

<sup>4</sup> (i) **Alle** Schneemänner haben einen Besen – Die Schneemänner haben **alle** einen Besen  
all snowmen have a broom the snowmen have *all* a broom  
(ii) **Jeder** Schneemann hat einen Besen – Die Schneemänner haben **jeder** einen Besen  
each snowman has a broom the snowmen have *each* a broom

this possibility is not given, children are believed to be puzzled as to why the experimenter wants to know whether a statement is true.

Crain et al. (1996) and Gualmini et al. (2003) insist on the children's full *semantic* competence for quantifiers and locate the source of errors in the test materials and the test situation. According to them, children do not lack knowledge of any aspect of quantification. The kids' incorrect responses are errors due to flaws in the experimental design. Other pragmatic sources of the errors are sought in the children's limited capacities of identifying the domain restrictions of indefinites as existentially quantified expressions (Rakhlin 2007: 246) relative to varying external contexts.

In several experiments, Brooks and associates have provided specific conditions for plausible dissent but the answer patterns did not change (Brooks and Braine 1996; Brooks et al. 2001). On the other hand, Sugisaki and Isobe (2001) explicitly tested against "plausible dissent" as the crucial factor and demonstrated that children expected to fail show adult-like performance even in situations in which the condition of plausible dissent is not satisfied.<sup>5</sup> Geurts (2003) objects, too, and drily points out that in these accounts the main point is missed:

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 [1996] lights, is hopelessly flawed? Crain et al. have very little to say about this. (Geurts 2003: 204)

Roeper and de Villiers (1993) conjectured a *syntactic* deficiency behind the high failure rate. They argue that children fail to identify the quantifier as part of the noun phrase; instead it would be regarded as a sentential adverb with scope over the whole sentence, whence the preference for the wrong answer and exhaustively paired readings. In other words, "[Every boy] rides on an elephant" would be mistaken as "[Every (time)] boy rides on an elephant."

From the perspective of a V2-language like German, this kind of syntactic parsing failure can be ruled out right away. In German (and in all other Germanic languages, except English), the first position in a declarative clause is reserved for a *single* syntactic unit. Hence, *jedes Auto* / *alle Autos* 'every/each car' / 'all cars' is easily identified as a single constituent in a clause like [*Jedes Auto*] *steht* <sub>Vfin.</sub> *in einer Garage* 'every car sits in a garage' simply because it

---

<sup>5</sup> In the experiments to be reported below, the pragmatic felicity was procured as follows. The experimenter would tell the children that her set of pictures had previously gotten messed up. Each child is explicitly encouraged to assist the experimenter in deciding which picture should go to the "is wrong" pile, and which one to the "is correct" pile.

precedes the finite verb. In German, being a V2 language, only a *single* constituent may precede the finite verb in a non-embedded declarative. If *jeder* were regarded as a separable item, it would have to be directly followed by the finite verb. German-speaking children master the *verb-second constraint* at latest by the age of three (see Tracy 1991; Tracy and Thoma 2009), but their failure rates with universally quantified subjects are parallel to English-speaking children at the age of four and later. Hence, it seems to be safe to exclude that they misidentify *alle Autos* ‘all cars’ or *jedes Auto* ‘every/each car’ by something like ‘[[*always/in each case*] [[cars]...]]’.<sup>6</sup>

A *semantic* deficiency with the very same confound was the initial presumption. According to Philip (1995: 53), children initially tend to construe *universal* quantifiers like adverbials, quantifying over events rather than individuals. For whatever unknown reason, children would assign the logical form which older children and adults assign to a sentence containing an adverb of quantification to a sentence with a universally quantified subject.

This characterization introduces a central premise without independent evidence. Geurts (2003) emphasizes that “in accounts 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.” (Geurts 2003: 118)

Geurts (2003) elaborates Drozd’s (2001) hypothesis, who suspects that children are unable to handle the appropriate distinctions between universal quantifiers (strong quantifiers) and existential quantification (weak quantifiers). Geurts (2003: Figure 1) specifically locates the complications in the mapping of syntactic form to semantic representation. The children are well aware of the meaning of *all*, *each*, and *every* in isolation. They fail when it comes to produce the appropriate compositional semantics of a universally quantified subject (that is, a universal quantifier and its restrictor) and a predicate that contains a quantified object.

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. (Geurts 2003: 209).

---

<sup>6</sup> The debate about the scope of *nur* ‘only’ and *sogar* ‘even’ and V2 vs. V3 (cf. Büring and Hartmann 2001; Reis 2005; Meyer and Sauerland 2009) might be of interest in this context as well, as one of the reviewers suggests.

The terms “strong” and “weak” refer to the strong-weak distinction introduced by Milsark (1977).<sup>7</sup> Geurts’ statement seems to be too strong. If children “interpret all quantifiers as if they were weak”, this would include definite noun phrases, too, since they qualify as “strong”; see Barwise and Cooper (1981: 182).<sup>8</sup> In this paper, we shall limit our attention to universally vs. existentially quantified noun phrase and their mapping relations between their syntactic and their semantic representation, with set-inclusion vs. set-intersection as a crucial factor.

There are still other properties of the stimulus material that have been suspected to additionally facilitate errors. Freeman et al. (1982) found that *relative salience* tends to bias children’s responses one way or the other, a finding that was confirmed in experiments by Drozd and van Loosbroek (1999, 2006). The clause final position, for instance, is the position for a salient element because it is the position for the nuclear accent. So, in a stimulus such as “*Every car is in a garage*” the clause-final position of *in a garage* could be a potential source of misinterpretations.

### 3 Locating the source of the difficulties

Why should a weak quantifier be easier to handle? The crucial difference seems to lie in the task of computing the relation between the two sets involved in the extensional interpretation of the two quantifiers, that is, the restrictor set of the subject and the set that corresponds to the quantifier in the scope of the quantified subject. As Geurts (2003) emphasizes, a weak quantifier involves a *set-intersection* relation, whereas a strong quantifier calls for a *set-inclusion* relation. In logical terms, this is the difference between the logical form of a *conditional* (2a) and that of a *conjunction* (2b).

- (2) a. [All As are Bs]  $\equiv$  For every x [**if** (x is an A) **then** (x is a B)]
- b. [Some As are Bs]  $\equiv$  There are *some* x such that [(x is an A) **and** (x is a B)]
- (3) a. ‘All As are Bs’ implies:  $(A \cap B) = A$ , that is,  $A \subseteq B$ ,
- b. ‘Some As are Bs’ implies:  $(A \cap B) \neq A$  (if A is not properly included in B)

---

<sup>7</sup> Quantifiers which can appear in a context such as (i), are called “weak” quantifiers; those that cannot are called “strong”.

(i) *There is/are ---- unicorn/uniorns in the garden.*

According to this operational definition, indefinites are weak while definites are strong.

<sup>8</sup> This point has been raised by one of the reviewers.

In set-theoretic terms, the restrictor set of a *universally* quantified noun phrase (3a) is a *subset* of the predicate set (set inclusion). In other words, the members of the A set are a (proper) subset of the B set. On the other hand, in a so-called weakly quantified noun phrase denotations (3b), that is, in sentences with existentially quantified subjects, the restrictor set *intersects* with (the extension of) the predicate (viz. set intersection) in their extensional interpretation. Moreover, the conjunction relation (which corresponds to set intersection in [2b]) is *commutative* while the implication relation (which corresponds to set inclusion in [2a]) is not. In sum, it is plausible that this poses problems for children when they are confronted with the scopal interaction of quantifiers in a clause.

In this situation – a bunch of competing hypotheses that seek the source of the difficulties in entirely different areas, ranging from test artefacts to genuine semantic acquisition issues – it occurred to us that a relevant question had not been asked or investigated before. Here it is: What will happen if a typical stimulus set that has been employed in previous investigations is replaced by a stimulus set with *semantically equivalent* stimuli that contain a (negated) existentially quantified subject rather than a universally quantified one? In other words, in addition to the standardly used stimuli such as (4a), we would also test a corresponding set of stimuli such as (4b).

- (4) a. **Every** boy walks with a balloon  
       b. **No** boy walks **without** a balloon

The idea amounts to manipulating the type of the quantifier while keeping everything else equal, as far as feasible.

## 4 The idea to be tested

If Geurt's (2003) hypothesis on the semantic source of the children's difficulties is basically correct and the crucial factor has to be located in the semantic nature of the quantifier indeed, and in particular in the difficulties posed by universally quantified subjects having an existentially quantified item in their scope, manipulating this factor should produce immediate effects in the test results. In the typically employed picture selection tasks, (4a) and (4b) are expected to elicit the same choice of pictures for a correct match (or an erroneous mismatch). However, since (4a) is universally quantified, the semantic processing of this quantifier is predicted to trigger more mistakes than the processing of the weakly quantified subject in (4b).

If on the other hand, the source of the errors with structures like (4a) was mainly a pragmatic problem of the children with the specific tests (see Section 2, on plausible denial or dissent), stimuli such as (4b) should trigger roughly the same amount of errors as stimuli of the type (4a). After all, they are communicatively equivalent and match with the same pictures in the selection task.

It is a logical truth that every universally quantified expression can be converted into a truth-conditionally *equivalent* existentially quantified expression. In other words, universal and existential quantification can be inter-defined by means of negation (5). For any sentence with a universally quantified subject there is a semantically equivalent sentence with a negated indefinite subject combined with a negated predicate.

$$(5) \quad \forall x P(x) \equiv \neg \exists x \neg P(x)$$

A pair of sentences such as (6a,c) is a linguistic instantiation of the equivalence law in (5). In the experiment, stimuli such as (6a) were tested against stimuli such as (6b). The stimulus type (6b) situates the negation of the predicate in the preposition *ohne* ‘without = not with’.<sup>9</sup>

- (6) a. **Jedes** Mädchen ist mit einer Katze da  
       no girl is with a cat here  
       ‘Every girl is here with a cat’
- b. **Kein** Mädchen ist ohne eine Katze da  
       no girl is without a cat here
- c. **Kein** Mädchen ist nicht mit einer Katze da  
       no girl is not without a cat here

We explicitly did not consider testing stimuli with an explicit negation particle such as in (6c) for two reasons. First, avoiding explicit double negation by *kein* ‘no’ in combination with *nicht* ‘not’ would avoid a processing difficulty (Just and Carpenter 1971). Second, and even more importantly, we had to eliminate the risk that children misinterpret a doubly negated sentence like (6c) as a

<sup>9</sup> Here are two examples that illustrate the substitutability of *nicht mit* ‘not with’ by *ohne* ‘without’.

- (i) Er ist **nicht mit** ihr verreist – Er ist **ohne** sie verreist.  
     he has not with her travelled he has without her travelled
- (ii) er wurde **nie mit** ihr gesehen – Er wurde **stets ohne** sie gesehen.  
     he was never with her seen he was always without her seen



case of negative concord. After all, all Bavarian-Austrian dialects use negative concord.<sup>10</sup> The children we tested grow up in a dialect area with negative concord. So we would not have been able to exclude potential interferences. In the negative concord interpretation, (6c) would not be a paraphrase of (6a). It would be the negation of (6a), meaning ‘No girl was here with a cat’.

(6b), on the other hand, captures the equivalence relation in (4) since the preposition *ohne* ‘without’ allows to unambiguously (though implicitly) negate the phrase that contains the existentially quantified predicate and avoid the aforementioned problems.

The stimuli used in the experiments are listed in the Appendix. They minimally differ from (6b). The stimulus set used in all test sessions was unintentionally streamlined with colloquial usage. In the *without*-phrase, bare nouns were used instead of nouns with an indefinite article. For instance, *ohne Geschenk* ‘without present’ would be used in place of *ohne ein Geschenk* ‘without a present’. In the particular context of the picture matching task, the variant without an indefinite article would match or mismatch with the same pictures that the variant with an indefinite article would match or mismatch.

In sum, the issue to be empirically clarified is this: How do children perform on patterns like (6b) in comparison to patterns like (6a)? If it turns out that patterns such as (6b) are mastered better than patterns like (6a), even despite two instances of negation being involved, then this would lend support to Geurt’s (2003) hypothesis of a semantic source of the difficulties displayed by the children’s reactions rather than sources such as pragmatics or test artefacts. (6a) and (6b) are exchangeable in terms of pragmatics since they express the same proposition, but they express the same proposition by different combinations of semantic structures.

A final remark is due on the acquisition of negation, since the existentially quantified equivalents of universally quantified sentences crucially involve two negated forms. We saw no reason to doubt what has been noted in the literature on the acquisition of negation (e.g. Clahsen 1988; Wojtecka et al. 2011). Children at the age of the children that were tested master negated sentences in either form, that is negation by a negation particle (*nicht* ‘not’), by a negated determiner (*kein* ‘no’), or by a negative preposition (*ohne* ‘without’). Wojtecka et al.

---

<sup>10</sup> Here is an example of negative concord: *I hob koa Bia ned drungga* (lit. I have no beer not drunk; ‘I did not drink a beer’). The Standard German rendering is: *Ich habe kein Bier getrunken* (lit. I have no beer drunk). Negative concord marks every indefinite by its negated form but the utterance counts as simply negated: *Bei uns hod no **nia** koana Hunga **need** leidn miassa* (lit. With us has still never nobody of hunger not suffer must; ‘With us, nobody has ever had to suffer from hunger’) (see Merkle 1975: 155–156).

(2011) report 90 % success rate for identifying false negated clauses for the age group 4;0 to 4;11. The participants we tested ranged from 4;04 to 5;05.

## 5 Test hypothesis

If Geurts' (2003) hypothesis is right, that is, if the source of the problem is a semantic one, namely a problem of dealing with universally quantified subjects in relation to a quantified item in its scope, then we have to predict a significant difference. Replacing universal quantification by the truth-conditionally equivalent negated existential quantification is expected to have a positive effect on the mastering of the task since this amounts to replacing a universally quantified subject (strong quantifier) by an existentially quantified one (weak quantifier) and thereby removing the suspected obstacle for the successful processing of these structures by children in the given age range.

The assumption we start from has already been introduced in the preceding section. We adopt and test a prediction of Geurts' hypothesis. This hypothesis states that the children's difficulties with the picture matching reflect difficulties of mapping syntactic form to semantic representation in the case of universal quantification. The relation between A and B in (7a) is intersective. In (7b), it is not (Geurts 2003: 210).

- (7) a. Some A are B Semantic representation:  $\langle \text{some} \rangle [x: A(x), B(x)]$   
 b. All A are B Semantic representation:  $[x: A(x)] \langle \text{all} \rangle [x: B(x)]$

When judging the overall complexity of (6a), (6b), one will note that (6a) and (6b) involve two independent complicating factors, respectively. Mastering (6a) presupposes the mastering of a semantic form such as (7b) which involves an *inclusion* relation as opposed to the conjunction in (7a). The mastering of (6b) should be easier since the conjunction in the logical representation triggers an intersection relation which is the canonical relation that kids apply for combining a subject with a restrictor and the predicate. In Geurts' (2003: 200) words, "weak quantifiers can be handled with simpler means than strong ones." The simple means for (7a) is set intersection. For (7b), however, the relation between the set of the restrictor and the predicate set must be construed as an *inclusion* relation, that is a subset relation rather than a set *intersection*.

On the other hand, (6b) necessitates the simultaneous mastering of *two* occurrences of negation. One of the two negations, viz. the negation of the existential, is explicitly coded by *kein* 'no' while the second negation is provid-

ed by the preposition *ohne* ‘without’. ‘Without x’ means ‘not with x’ or ‘with no x’. When judging the test results one should bear in mind that this is a complicating factor. If, nevertheless, the type (6b) turns out to be significantly easier than (6a), then this strengthens the support for the hypothesis that the type of quantifier is the essential source of the difficulties.

- (8) **Set-inclusion-vs.-intersection hypothesis:** Stimuli with a *universally quantified subject* provoke a higher amount of deviant answers than *truth-conditionally equivalent* stimuli with a *negated existentially quantified subject* plus an (implicitly) negated object. The correct understanding of universally quantified subjects rests on the understanding of an entailment relation<sup>11</sup> (which amounts to a *set-inclusion* relation), while existential quantification is based on a conjunction relation (which amounts to a *set-intersection* relation).

A basic developmental step in the children’s competence at the crucial age seems to be the ability to handle universally quantified subjects in relation to quantified items in their scope by constructing the proper semantic forms and their denotations. For children, this requires enlarging the tool kit by *set inclusion* (in correspondence to the conditional in the predicate logical form of universal quantification) in addition to the combinatorial strategy of *set intersection* (based on conjunction in the logical form).

The proponents of pragmatic deficits as a source of mistakes, as discussed above, are bound to predict that (6a) and (6b) will pose the very same pragmatic difficulties. The stimuli types (6a) and (6b) are semantically equivalent, that is, they are verbalizations of truth-conditionally equivalent propositions. Whatever context “defect” in terms of “(im)plausible dissent” or “(im)plausible denial” holds for (6a) will equally apply to (6b).

On the other hand, if “strong” vs. “weak” quantification is indeed at the root of the problem as Drozd (2001) or Geurts (2003) assume (with different implementations, though), the prediction is that the pattern (6b) will be mastered more easily than the pattern (6a).

---

<sup>11</sup> Every A is B = For every x, if x is A then x is B. So, the A set is a *subset* of the B set, that is, *included* in B.

Some A are B = For some x, x is A *and* x is B. So, A and B *intersect*.

## 6 Test subjects, procedure, stimuli, operationalization

**Test subjects:** The tests were carried out in two kindergartens, one in the city of Salzburg and one in the neighboring town Kuchl. Sixteen randomly chosen children were tested, with the following age distribution: mean age 4;09 (youngest participant: 4;04; oldest one 5;05). The results of *one* child had to be excluded because the child quit during the testing session. Each child was tested individually by a female experimenter in a quiet room. Testing a child took 10–15 minutes. During the testing, the child received only encouraging feedback from the experimenter independent of the correctness of the response.

**Procedure:** In order to create a pragmatically natural situation, each child was first told by the female experimenter that her set of pictures had gotten in disorder and that she asks the child for assistance when putting the pictures in order again, as correct or incorrect. Next, the experimenter would read the respective stimulus sentence, show a picture and ask the child “*Stimmt das?*” ‘Is this correct?’. The stimuli were presented in a pseudo-randomized order.

**Stimuli:** The graphic stimuli were professionally designed, depicting a situation described by the stimulus sentence. For each linguistic stimulus, there were two pictures, depicting situation 1 and situation 2, as indicate in (1b) and (1c), respectively. Each linguistic stimulus was tested twice, in random order, with a picture for situation 1 and a second time with a picture for situation 2.

**Operationalizing of the testing:** 16 pictures; eight for situation (1b), eight for situation (1c). Each picture was tested in random order with one of the two types of stimulus sentences, that is, with a universal quantifier (9a), (9c) and with a negated existential plus ‘without’ (9b). In addition, a third linguistic stimulus type was tested, as a control for the children’s correct understanding of the pictures. The third type contained a paraphrase<sup>12</sup> of the negated predicate in stimuli of the type (1b). For instance, ‘not in a tree’ was paraphrased by ‘in the air’ (9d). Unlike (9b), such a stimulus does not contain an implicit negation.

- (9) a. universal Q: *Jeder Hund hat einen Knochen* ‘every dog has a bone’  
 b. neg. exist. Q + ‘without’: *Kein Hund ist ohne Knochen* ‘no dog is without bone’

---

<sup>12</sup> “Paraphrase” in this case means that the predicate of these stimulus sentences denotes a consequence of the negated predicate of the universally quantified stimulus sentence, e.g. “Every car is in a garage” – “No car is outdoor”.

- c. universal Q: *Jeder Vogel sitzt auf einem Baum* ‘every bird sits in a tree’
- d. neg. exist. Q + paraphrase: *Kein Vogel ist in der Luft* ‘no bird is in the air’<sup>13</sup>

The set (9d) primarily served as a control for the understanding of negated existential quantification in the subject as well as for the pragmatic adequacy. Since stimuli like (9d) have predicates that denote the opposite of the predicates in (9a), a defect with respect to “plausible denial” would be likely to show with this kind of stimuli as well.

## 7 Results

The experiment produced results that support the test hypothesis. If the results are representative, they indicate that the tested variable (viz. semantically equivalent stimuli with either a universally quantified subject or a negated existential as subject) accounts for a very large portion of the errors found in the picture matching tasks. As this study has pilot character (limited number of stimuli, limited number of test subjects), we regard the findings as provisional.

The experiment tested three conditions, namely (i) the responses to stimuli with universally quantified subjects, (ii) with negated existentially quantified subjects plus the preposition ‘without’, and (iii) with negated existentially quantified subjects plus paraphrased predicate (lexical converses).

Situation (1) refers to the extra object of situation (1b) in Section 1; situation (2) is the extra agent situation of (1c) in Section 1. The raw data are presented in Table A in the Appendix and in Schörghofer-Essl (2012). As this is a pilot study, the number of subjects and stimuli are not of the size that would warrant detailed statistical analyses. Instead, we present merely the results (plus raw data), with a minimum of descriptive statistics.

### Condition (i) – Universal quantifier

Situation (1): **23 %** correct; Possible max. score: 8;  $M = 1.87$ ;  $Mdn = 1.0$ ;  $SD = 1.41$   
 Situation (2): **58 %** correct; Possible max. score: 8;  $M = 4.60$ ;  $Mdn = 5.0$ ;  $SD = 2.10$

---

<sup>13</sup> The chosen converse predicated is relative to the pictures presented. In the case of this very stimulus, the picture showed only birds sitting in trees or flying in the air. Hence, ‘in the air’ is the converse of ‘in trees’ for this picture.

**Condition (ii)** – Negated existential quantifier with implicitly negated predicate  
 Situation (1): **48 %** correct; Possible max. score: 4;  $M = 1.87$ ;  $Mdn = 2.0$ ;  $SD = 1.46$   
 Situation (2): **70 %** correct; Possible max. score: 4;  $M = 2.80$ ;  $Mdn = 3.0$ ;  $SD = 0.94$

**Condition (iii)** – Negated existential quantifier and paraphrased predicate  
 Situation (1): **97 %** correct; Possible max. score: 4;  $M = 3.87$ ;  $Mdn = 4.0$ ;  $SD = 0.35$   
 Situation (2): **93 %** correct; Possible max. score: 4;  $M = 3.73$ ;  $Mdn = 4.0$ ;  $SD = 0.46$

In general, these results confirm the initially formulated expectation that children master negated existential quantifiers in combination with a negated predicate much better than the truth-conditionally equivalent universal quantifiers combined with the predicate.

For the crucial items of Situation (1), the mean percentage of correct answers under condition (ii) is more than 100 % higher than under condition (i). In other words, the percentage of errors triggered by stimuli with a universally quantified subjects is twice as high as the percentage of errors triggered by stimuli with a negated existential as subject.

As for an in-between-subjects comparison, the results are clear-cut, too. There is *no child* who masters more than 50 % of the universally quantified stimuli in the Type (1b) task correctly. On the other hand, almost half of children, namely seven out of 15, master between 75 % and 100 % of the tasks with a negated existential correctly. This can be verified by inspecting the table with raw data in the Appendix.<sup>14</sup>

## 8 Discussion

The success rate for the stimuli with a negated existential plus ‘without’ is twice as high as the rate for the corresponding stimuli with universal quantification. This amounts to a rather strong indication that hypothesis (8) is relevant, that is, that the crucial causal factor reflected in the children’s reaction patterns is the handling of set inclusion and set intersection in the handling of weak vs. strong quantifiers.

---

**14** When checking the raw data (Table A, Appendix), please be aware that the number of stimuli for the first type of task (universal quantification) was twice as many as for the following two tasks, that is eight, four, and four. So “four” in the first condition would be a 50 % success, while the same number of correct responses in the second and third condition would mean 100 %.

The answers to the quasi-control condition (iii) turned out as close to perfect. This indicates that the children understood the task and the stimuli. There are no indications of potentially interfering effects of test artefacts.

Table 1 presents the results for the eight contrasting stimulus pairs. In each of the four boxes in the left column, the first line is the result for the stimulus with universal quantification tested with a type-A picture (“Bild A”) presenting a situation of the type (1b) and the second line contains the type-B picture result (“Bild B”), presenting a (1b) situation. The first and the third, as well as the second and the fourth line in each box represent the results for the very same picture tested, either with the universal quantifier (line 1 and 3) or the negated existential quantifier (line 2 and 4). The figures under “correct” and “wrong” are the raw numbers (15 children).

**Table 1:** Stimulus sentences and correct/wrong reactions in comparison.

Stimulus sentences	Picture		Answers		Correct answers in percent
	Type <sup>15</sup>	Order <sup>16</sup>	correct	wrong	
<i>Jedes Mädchen streichelt eine Katze.</i>	Bild A	Bild 30	3	12	20.00%
<i>Jedes Mädchen streichelt eine Katze.</i>	Bild B	Bild 13	9	6	60.00%
<i>Kein Mädchen ist ohne Katze.</i>	Bild A	Bild 2	4	11	26.67%
<i>Kein Mädchen ist ohne Katze.</i>	Bild B	Bild 18	10	5	66.67%
<i>Jeder Hund frisst einen Knochen.</i>	Bild A	Bild 19	3	12	20.00%
<i>Jeder Hund frisst einen Knochen.</i>	Bild B	Bild 5	9	6	60.00%
<i>Kein Hund ist ohne Knochen.</i>	Bild A	Bild 27	7	8	46.67%
<i>Kein Hund ist ohne Knochen.</i>	Bild B	Bild 12	12	3	80.00%
<i>Jeder Mann bringt ein Geschenk.</i>	Bild A	Bild 28	4	11	26.67%
<i>Jeder Mann bringt ein Geschenk.</i>	Bild B	Bild 16	8	7	53.33%
<i>Kein Mann kommt ohne Geschenk.</i>	Bild A	Bild 7	9	6	60.00%
<i>Kein Mann kommt ohne Geschenk.</i>	Bild B	Bild 20	9	6	60.00%
<i>Jedes Pferd trägt einen Sattel.</i>	Bild A	Bild 23	3	12	20.00%
<i>Jedes Pferd trägt einen Sattel.</i>	Bild B	Bild 8	10	5	66.67%
<i>Kein Pferd ist ohne Sattel.</i>	Bild A	Bild 14	8	7	53.33%
<i>Kein Pferd ist ohne Sattel.</i>	Bild B	Bild 31	11	4	73.33%

**15** A stimulus picture of the type “Bild A” depicts the situation described in (1b) (see the introduction section), as illustrated by Figure 1. “Bild B” is a stimulus picture that depicts a situation described in (1c), as illustrated by Figure 2 in the introduction section.

**16** The “picture number” refers to the randomized order of presentation. For example “Bild 30” (picture 30), as in the first line, means that the picture was presented (again) after twenty nine picture cards had been already presented in the experiment (testing for type A or B effects).

The general tendency is robust and easy to grasp. There is a clear advantage for the (negated) existential condition. Except for the first box with only a 33% advantage, all the other results in the third lines (viz. negated existential) are at least 50% above the correct answers for the universal quantification condition.

Therefore, we conclude that the hypothesis we put to test – set inclusion vs. set intersection – is very likely a decisive factor for the failures with universally quantified stimuli in situations of the type (1b), Section 1. These findings are findings of a pilot study, with a restricted number of test subjects, and a restricted number of stimuli, of course.

The results match our expectations that children at the given age span have yet to find out that universal quantification necessitates the contemplation of set inclusion rather than set intersection. The data suggest that it takes some time for children to correctly differentiate between their preferred mode of set *intersection* in the building of semantic interpretations and the required mode of set *inclusion* for the restrictors of universal quantification.

Finally, we have to concede that the factor we investigated does not account for the whole amount of errors. Even if the percentage of correct answers doubles for the negated existentials in comparison to the universally quantified stimuli, there is still a gap between child and adult responses, which remains unaccounted for if the data reported here are representative.

This study has been planned as a pilot study, whence the proportionately small number of test subjects and test items, which do not warrant more fine-grained statistical analyses. However, the fairly clear-cut results encourage us to believe that a full-fledged study will support and reproduce the initial picture obtained by this exploratory study.

## 9 Summary

The results lend support to the test hypothesis. A high proportion of the failure rate of children in the well-known “quantifier spreading” experiments seems to be a genuine developmental effect in the acquisition of compositional semantics. We tried to verify this by testing both with universally quantified stimuli and with their logical equivalents in the form of negated existential quantification, with the following prediction: Children confronted with the same pictures will produce significantly more correct results with negated existentials than with the logically equivalent stimuli with universally quantified subjects. The results support this prediction.

The factor *set inclusion* vs. *set intersection* in computing the relations between a quantified subject and a quantified objects seems to have a strong



effect on the success rate of children of about four to five years of age when confronted with decision tasks on truth-conditionally equivalent sentences with universally quantified vs. negated existentially quantified subjects (with negated predicates). The (widely observed low) success rate for universal quantifiers *improves* by nearly 100% when the stimulus is replaced by a semantically equivalent stimulus with an intersective (= weak) quantifier (and a negated predicate). Nevertheless, there still remains a substantive proportion in the failure rates that distinguishes children of this age from older children or adults that is not covered by the factor investigated here. Moreover, the source of the higher percentage of false negative (high) and the generally lower percentage of false positive reactions is still not fully accounted for.<sup>17</sup>

If the findings of this pilot study are representative, they confirm the hypothesis that a major source of the particular difficulties for children in the given age group is the appropriate semantic processing of a *universally* quantified subject in combination with a quantifier in the predicate. This eliminates competing accounts based on pragmatic factors, test artefacts or the like, since these results are based on a precisely identified difficulty, namely the difficulty of mapping universally quantified subjects on the appropriate interpretations.

These difficulties are predicted but do not show when the linguistic form of a stimulus is changed while its semantic content remains the same, that is, if a stimulus with a universally quantified subject is replaced by a semantically and pragmatically equivalent stimulus with a negated, existentially quantified subject. So, pragmatic factors or any other non-linguistic factors would evidently not succeed in capturing this difference.

**Acknowledgments:** The division of labor for this paper was as follows: The experiments were carried out and analyzed by the second and third author. The first author proposed the research strategy, was the supervisor of the projects, and put this paper in writing. The paper is based on two related studies. The second author studied the effect of substituting a universal quantifier by a negated existential quantifier. The third author investigated the effects of the linguistic rendering of universal quantification by *alle* ‘all.PL’ vs. *jeder* ‘each.SC’, in proximate or distant linearization.

We are very grateful to Lilli Schörghofer-Essl for providing the graphic stimuli, to Ms. Christina Brandauer (Kindergarten Kuchl) and Ms. Irene Mellmer

---

<sup>17</sup> Given that the mastery of false negatives precedes the mastery of true negatives (see Wojtecka et al. 2011), an analogous effect may play a role for the phenomena under discussion. In one case (situation 1b), children falsely *deny*, whereas in the other case (situation 1c), they falsely *accept* that a given description is correct. At least part of the different failure rates could find an explanation along this line.

(Pfarrkindergarten Herrnau) for their kind authorization of the testings, and last but not least to three anonymous reviewers, whose suggestions and criticism greatly improved the paper. Remaining shortcomings are to be blamed on the first author, of course.

## References

- Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2). 159–219.
- Brooks, Patricia. J. & Martin D. S. Braine. 1996. What do children know about universal quantifiers ‘all’ and ‘each’? *Cognition* 60(3). 235–268.
- Brooks, Patricia J. & Irina Sekerina. 2006. Shortcuts to quantifier interpretation in children and adults. *Language Acquisition* 13(3). 177–206.
- Brooks, Patricia. J., Martin D. S. Braine, Gisela Jia & Maria da Graca Dias. 2001. Early representations for ‘all’, ‘each’, and their counterparts in Mandarin Chinese and Portuguese. In Melissa Bowerman & Stephen C. Levinson (eds.), *Language acquisition and conceptual development*, 316–339. Cambridge: Cambridge University Press.
- Büring, Daniel & Katharina Hartmann. 2001. The syntax and semantics of focus-sensitive particles in German. *Natural Language and Linguistic Theory* 19(2). 229–281.
- Clahsen, Harald. 1988. Kritische Phasen der Grammatikentwicklung. Eine Untersuchung zum Negationserwerb bei Kindern und Erwachsenen. *Zeitschrift für Sprachwissenschaft* 7(1). 3–31.
- Crain, Stephen, Rosalind Thornton, Carole Boster, Laura Conway, Diane Lillo-Martin & Elaine Woodams. 1996. Quantification without qualification. *Language Acquisition* 5(2). 83–153.
- Droz, Kenneth F. 2001. Children’s weak interpretations of universally quantified questions. In Melissa Bowerman & S. C. Levinson (eds.), *Language acquisition and conceptual development*, 340–376. Cambridge: Cambridge University Press.
- Droz, Kenneth F. & Erik van Loosbroek. 1999. Weak quantification, plausible dissent, and the development of children’s pragmatic knowledge. In Annabel Greenhill, Heather Littlefield & Cheryl Tan (eds.), *Proceedings of the 23<sup>rd</sup> Annual Boston University Conference on Language Development (BUCLD 23)*, 184–195. Somerville, MA: Cascadia Press.
- Droz, Kenneth F. & Erik van Loosbroek. 2006. The effect of context on children’s interpretation of universally quantified sentences. In Veerle van Geenhoven (ed.), *Semantics in acquisition*, 115–140. New York: Springer.
- Freeman, Norman H., Chris G. Sinha & Jacqueline A. Stedmon. 1982. All the cars – which cars? From word meaning to discourse analysis. In Michael Beveridge (ed.), *Children thinking through language*, 52–74. London: Edward Arnold.
- Freeman, Norman H. & Jacqueline A. Stedmon. 1986. How children deal with natural language quantification. In Ida Kurcz, Grace W. Shugar & Joseph H. Danks (eds.), *Knowledge and language*, 21–48. Amsterdam: Elsevier North-Holland.
- Geurts, Bart. 2003. Quantifying kids. *Language Acquisition* 11(4). 197–218.
- Gualmini, Andrea, Luisa Meroni & Stephen Crain. 2003. An asymmetric universal in child language. In Matthias Weisgerber (ed.), *Proceedings of Sinn und Bedeutung 7*

- (Arbeitspapiere des Fachbereichs Sprachwissenschaft, Nr. 114), 136–148. Konstanz: University of Konstanz.
- Inhelder, Bärbel & Jean Piaget. 1958. *The growth of logical thinking from childhood to adolescence*. London: Routledge and Paul Kegan.
- Just, Marcel Adam & Patricia Ann Carpenter. 1971. Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior* 10(3). 244–253.
- Merkle, Ludwig. 1975. *Bairische Grammatik*. München: Hugendubel.
- Meyer, Marie-Christine & Uli Sauerland. 2009. A pragmatic constraint on ambiguity detection: A rejoinder to Buring and Hartmann and to Reis. *Natural Language and Linguistic Theory* 27(1). 139–150.
- Milsark, Gary L. 1977. Toward an explanation of certain peculiarities in the existential construction in English. *Linguistic Analysis* 3. 1–30.
- Philip, William Churchill Houston. 1995. *Event quantification in the acquisition of universal quantification*. Amherst, MA: University of Massachusetts dissertation.
- Rakhlin, Natalia. 2007. A new pragmatic account of quantifier-spreading. *Nanzan Linguistics* [Special Issue], 1. 239–282.
- Reis, Marga. 2005. On the syntax of so-called focus particles in German: A reply to Buring and Hartmann 2001. *Natural Language and Linguistic Theory* 23(2). 459–483.
- Roeper, Tom & Jill de Villiers. 1993. The emergence of bound variable structures. In Eric Reuland & Werner Abraham (eds.), *Knowledge and language*, vol. 1, *From Orwell's problem to Plato's problem*, 105–139. Dordrecht: Kluwer.
- Schörghofer-Essl, Christina. 2012. Interpretation von Quantoren im Kindergartenalter. Eine empirische Studie zu „Quantifier Spreading“. Salzburg: University of Salzburg MSc.Thesis. [http://www.uni-salzburg.at/fileadmin/oracle\\_file\\_imports/1859246.PDF](http://www.uni-salzburg.at/fileadmin/oracle_file_imports/1859246.PDF) (25 January 2017).
- Seethaler, Karin. 2012. Steht jedes Auto in einer Garage? – Zum Umgang von Vorschülern mit den Universalquantoren „alle“ und „jeder“. Salzburg: University of Salzburg MSc. Thesis. [http://www.uni-salzburg.at/fileadmin/oracle\\_file\\_imports/2095309.PDF](http://www.uni-salzburg.at/fileadmin/oracle_file_imports/2095309.PDF) (25 January 2017).
- Smith, Carol L. 1979. Children's understanding of natural language hierarchies. *Journal of Experimental Child Psychology* 27(3). 437–458.
- Smith, Carol L. 1980. Quantifiers and question answering in young children. *Journal of Experimental Child Psychology* 30(2). 191–205.
- Sugisaki, Koji & Miwa Isobe. 2001. Quantification without qualification without plausible dissent. *Proceeding of the SULA Conference (Semantics of Under-Represented Languages)*. Amherst, MA: University of Massachusetts. [http://faculty.human.mie-u.ac.jp/~sugisaki/Papers/Sugisaki\\_Isobe\\_2001\\_SULA.pdf](http://faculty.human.mie-u.ac.jp/~sugisaki/Papers/Sugisaki_Isobe_2001_SULA.pdf) (25 January 2017).
- Tracy, Rosemarie. 1991. *Sprachliche Strukturentwicklung. Linguistische und kognitionspsychologische Aspekte einer Theorie des Erstspracherwerbs*. Tübingen: Narr.
- Tracy, Rosemarie & Dieter Thoma. 2009. Convergence on finite V2 clauses in L1, bilingual L1 and early L2 acquisition. In Christine Dimroth & Peter Jordens (eds.), *Functional categories in learner language*, 1–43. Berlin & New York: Mouton de Gruyter.
- Wojtecka, Magdalena, Corinna Koch, Angela Grimm & Petra Schulz. 2011. Production and comprehension of sentence negation in child German. In Angela Grimm, Anja Müller, Cornelia Hamann & Esther Ruigendijk (eds.), *Production and comprehension asymmetries in child language*, 217–245. Berlin & New York: de Gruyter.

Appendix

Table A: Raw data of experiment 2 – correct answers.

Correct $\forall x$	Type (1b)	Type (1c)	Correct $\neg \exists x$	Type (1b)	Type (1c)	Corr. lex. con.	Type (1b)	Type (1c)
Child 1	4	5	Child 1	4	4	Child 1	4	4
Child 2	0	5	Child 2	0	3	Child 2	4	3
Child 3	1	6	Child 3	3	2	Child 3	4	3
Child 4	4	0	Child 4	0	4	Child 4	3	4
Child 5	3	2	Child 5	3	2	Child 5	4	4
Child 6	0	3	Child 6	2	2	Child 6	4	4
Child 7	2	6	Child 7	4	1	Child 7	4	3
Child 8	3	5	Child 8	2	2	Child 8	4	4
Child 9	1	5	Child 9	1	3	Child 9	4	4
Child 10	1	8	Child 10	2	4	Child 10	4	3
Child 11	1	5	Child 11	0	3	Child 11	4	4
Child 12	1	7	Child 12	1	4	Child 12	4	4
Child 13	4	2	Child 13	4	3	Child 13	3	4
Child 14	1	4	Child 14	1	3	Child 14	4	4
Child 15	2	6	Child 15	1	2	Child 15	4	4
mean	1.87	4.60		1.87	2.80		3.87	3.73
median	1.00	5.00		2.00	3.00		4.00	4.00
Sd	1.41	2.10		1.46	0.94		0.35	0.46
max score 8			max score 4			max score 4		
Abs.	28	69		28	42		58	56
% corr	23.33	57.50	% corr	46.67	70.00	% corr	96.67	93.33

Note that the *maximum score* a child can reach is *different* for the first set (stimuli with universal quantification). For this set of stimuli, the maximum score is 8. For the other two sets, that is, negated existential plus a ‘without’-phrase and negated existential plus a lexically converse predicate), the maximum score is 4 per set. So, in toto, each child was tested with 16 stimuli, that is, *eight plus four plus four*.

Because of the high inter-individual variation ( $Sd > 1,4$ ), the preferred indicator is not the mean but the median. It is twice as high for the negated existentials in comparison with the universally quantified stimuli.