(Quantifier) scope judgments
Kriszta Szendroi
UCL

1 Introduction

1.1 Quantifier scope interactions: a semantic ambiguity?

Sentences like (1a)-(4a) are arguably ambiguous in English. They contain quantifiers and operators whose interaction gives rise to the ambiguities illustrated in (1b)-(4b) and (1c)-(4c). The (b) examples illustrate what is usually called the *overt scope* reading, where the surface c-command relations between the quantifiers and operators are reflected in their scopal relations. This reading is also sometimes called 'isomorphic', again, because syntactic structure and semantic scope are matched. The (c) examples give the so-called *inverse scope* readings, where the surface c-command relations are reversed: here the scopal relations in meaning are the opposite of the c-command relations in surface syntax. The readings that involve sets of individuals for both quantifiers, rather than a single individual for the existential and a set of individuals for the universal, i.e. (1c) and (2b) are also sometimes called *distributive* readings.

(1) a. A doctor advised every nurse.
b. a > every: There is a doctor that advised every nurse.
c. every > a: For every nurse there is a doctor that advised her, but not necessarily the same doctor.

(2) a. Every nurse assisted a doctor.
b. every > a: Every nurse assisted a doctor, but not necessarily the same doctor.
c. a > every: There is a (specific) doctor and every nurse assisted him/her.

(3) a. All doors in this car will not open at the next station.[1]
b. all > not: All the doors in the carriage will not open at the next station.
c. not > all: It is not the case that all the doors in this carriage will open at the next station.

(4) a. The detective didn't find two guys.
b. not > two: It is not the case that the detective found two guys.
c. two > not: There are two guys that the detective did not find.

[1] Announcement on London's Hammersmith and City underground line before Baker Street station, in the final carriage, which is only partially aligned with the platform at that station. The announcement is presumably intended with the inverse scope reading as it continues with *Please use other doors*.

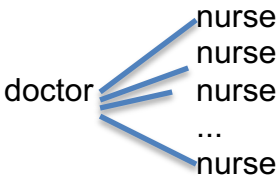
It has long been debated in the linguistic literature how best to represent such ambiguities in syntax and semantics, or in fact, whether to represent them at all. Ambiguities are present in other sentences, for instance, the sentence in (5) is recognised to be ambiguous between a reading where the PP modifier attaches to the NP and a reading where it modifies the VP.

(5) a. I see the man with the binoculars.
b. NP-modifier reading: the man has binoculars
c. VP-modifier reading: I see the man through binoculars

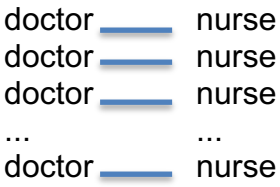
This ambiguity thus arises because of the different syntactic positioning of the PP modifier. There is in fact no doubt that the sentence is ambiguous in this way, as the two readings have distinct truth conditions, i.e. they are true in different situations.

But in sentences like (2), the readings are not so clearly distinct. In fact, one reading is logically entailed by the other: In any situation where a specific doctor is being assisted by every nurse, it is also true that every nurse assists a doctor. In other words, in any situation where (2c) is true, (2b) is also always true. This fact has been used to argue against the presence of a real semantic ambiguity in such sentences. In order to see how the argumentation goes let us consider the schematic situations in (6).

(6) a.



b.



On the face of it, (6b) depicts the inverse scope reading in (2c), while (6a) depicts the overt scope reading in (2b). But let us assume for argument's sake that (2a) is not actually ambiguous, rather the only reading that is available is the overt scope reading, where surface c-command relations correspond to the semantic scope of the quantifiers, i.e. (2b). This reading is true in both situations depicted in (6). So, the argument goes, there is no actual inverse scope reading of (2), rather the situation in (6a) is just a special case of the distributive type of situations-- one where the doctors assisted by the nurses happen to be a single individual. So, this argumentation can be used to argue that in semantics, (2a) only has one reading, the overt scope one, (2b).

The same issue arises with respect to (1a). Here, however, the more general situation (i.e. the one involving (potentially) different doctors, i.e. (6b))

corresponds to the reading that represents inverse scope, i.e. (1c). The overt scope reading, (1b), corresponds to the more specific situation, where the doctors who advise the nurses happens to be a single individual. So, if we want to treat these two situations be compatible with a single semantic meaning, as we did with sentences like (2), then we are forced to assign the inverse scope reading, i.e. (1c), as the only meaning of (1a). So, we are forced to abandon the parallelism between surface syntactic c-command relations and scopal order in the semantics.

But this is clearly counterintuitive. Why would two quantifiers take obligatorily overt scope in sentences like (2) but inverse scope in sentences like (1)? One could try to posit that the quantifier *every* always takes wide scope over an existential quantifier and thus both (1c) and (2b) would be derived, with (1b) and (2c) being special cases of them, respectively. This would be a good fix here, but this does not seem to hold generally. In sentences like (3), native speakers clearly easily access both the overt scope reading in (3b), with the universal taking wide scope, and also the inverse scope reading in (3c), where the negator takes scope over the universal. The same holds for (4).

In any case, (3b) and (3c), and (4b) and (4c) are not in entailment relations. The truth conditions of (3b) are orthogonal to the truth conditions of (3c): one can conjure up a situation that makes (3b) true and (3c) false, and vice versa. The same holds for (4b) and (4c). So, (3) and (4) are similar to the PP-modifier ambiguity sentence in (5). It is natural to assume then that just like (5), at least (3) and (4) would involve a proper semantic ambiguity. The only difference is that in (3) and (4) the ambiguity is not due to the syntactic position of a PP-modifier, but rather to the relative scopal relations of a quantifier and negation. But then it seems to lack generality to then not propose the same kind of scopal ambiguity to be present in sentences with two quantifiers like (1) and (2).

In addition, it is also the case that speakers have a psychological sense of ambiguity in examples like (1-4). What I mean by that is that the sentences are recognised to be ambiguous by speakers. Once they are encouraged to engage in metalinguistic considerations about the meaning of the sentences, they see them as ambiguous between the overt scope reading and the inverse scope reading. Speakers can offer judgments based on the distinct readings. Very often, like in the case of (1), they have strong preferences. Anyone who has ever taught the grammar of such sentences to first-year undergraduates can testify that many of them insist that they only allow the overt scope reading, at least until they are confronted with syntactically analogous[2] but lexically biased examples like Hirschbühler's (1982) famous sentence, in (7), where the inverse scope reading 'shines through':

(7) An American flag was hanging in front of every building.

[2] Near-analogous, to be precise, as such examples often have verbs with prepositional complements rather than NP direct object arguments.

It is also often claimed (see e.g. Jackendoff 1972), although empirically this has proven difficult to substantiate (see Syrett et al. 2014 for a successful attempt), that appropriate prosody can be used to disambiguate the two scopal readings:

(8) All the men didn't eat.
  a. falling tune: overt scope
  b. rising tune: inverse scope

In sum, eventually, a consensus was reached in the field that the ambiguities illustrated in examples (1-4) are semantically real and should be represented in our grammar of English.

1.2 Research questions

One of the first questions that then arises is whether the ambiguities should be represented in the syntax and if so, how. Recall the examples with PP modifiers like *with binoculars,* i.e. (5). These examples represent structural ambiguities, where the syntactic position of the PP determines which reading the sentence actually has. It is not immediately obvious what kind of structural account one could give to the scopal ambiguities in (1-4). Or whether the same analysis should be adopted for all four types of sentences. As we will see later, experimental work can be illuminative in this respect.

Another group of questions concerns how people understand such ambiguous sentences in real time. Going back again to the structural ambiguity in sentences like *I see the man with the binoculars*, it is well-known that people prefer, at least temporarily, while they are parsing the sentence, the reading where the PP modifies the whole VP. One may ask if there are similar preferences in the case of scopal ambiguity. Does the hearer commit to one of the readings early on during parsing, say to the overt scope reading? Or do hearers maintain the ambiguity as long as disambiguating evidence is encountered and only then commit to one of the readings? And what exactly counts for disambiguating evidence?

Another important, and potentially related, issue concerns the role of context. The role of context is especially relevant because it is clear that the two readings associated with sentences involving an existential and a universal are not equal in terms of conceptual representation. Recall that the distributive scope reading (i.e. 1c and 2b) involves, or at least potentially involves, a set of individuals taking part in a series of events, while the wide-scope existential reading necessitates the presence of only one such individual.

Crain & Steedman (1985) and Altmann & Steedman (1988) argued that the preferred interpretation adopted by the parser for any kind of ambiguity is the one that carries the fewest unsatisfied presuppositions. In other words, hearing a sentence like (1) out of the blue, with no previous discourse context, the parser would prefer to assign the overt scope interpretation because that requires accommodating the existence of a single doctor in the discourse context, while the inverse scope reading (at least potentially, and according to

Fodor (1982) preferentially) requires the accommodation of a set of nurses. But crucially, Crain & Steedman (1985) demonstrate that such preferences are dependent on discourse context. If the sentence is not parsed out of the blue, but rather embedded in a discourse context that already establishes the existence of a salient set of nurses, then the parser should no longer have a preference for the overt scope reading. Taking that line of logic one step further, if the context is designed in such a way that it is heavily biased towards a distributive interpretation, perhaps one would even expect an inverse-scope parsing preference.

Based on such considerations, one might ask if speakers interpretation is sensitive to the discourse context in which the sentences occur. Do people have less of a preference, for say, the overt scope interpretation in a discourse situation that strongly favours an inverse scope interpretation? Do hearers actually parse such ambiguous sentences differently in contexts that favour one of the readings than in null contexts? More generally, one might ask if there is an extra processing cost associated with one of the readings. If so, that would be relevant for our linguistic analysis of these ambiguities. If one of the readings is easier to compute than the other one, perhaps that is due to that reading being syntactically simpler than the other one.

Finally, do children parse such sentences the same way as adults? Are they aware of the ambiguities or can be made aware of them the same way as adults? Do they have a preference for one of the readings? Is it an even stronger preference than that of adults? Given that they have a developing grammar as well as a fewer processing resources, one might expect children to behave differently from adults under certain scenarios. Again, this kind of information would ultimately be useful for determining the correct linguistic analysis of such ambiguities.

The rest of this chapter will review the psycholinguistic literature concerning scopal ambiguities in search of answers to these questions. We will consider how speakers assign the different readings to these scopally ambiguous sentences. We will see if some sentences are treated differently from others, and we will consider the time course of the parsing process. We will investigate the potential role different discourse contexts have in influencing people's interpretation of the sentences and whether children are different from adults in any of these respects.

1.3 Roadmap

Recall that a crucial difference presents itself in that the two readings obtained by the different scopal orderings of an existential and a universal quantifier are not distinct, and that this issue does not arise in the case of sentences with a quantifier and negation. Although, as we have established above, both types contain real, semantic scopal ambiguity, nevertheless it turns out that speakers do not treat these two types of scopal ambiguities exactly the same way. This directly influences the linguistic analyses we should adopt for different types of scopally ambiguous sentences.

The (non)-distinctness of the readings also presents a methodological issue. Since, it is logically impossible for instance to create a situation where say (1b) (i.e. There is a doctor that advised every nurse) is true and (1c) (i.e. For every nurse there is a doctor that advised) is false, it is problematic to test such non-distinct readings in any experimental task dependent on the truth conditions of the sentence, such as the Truth-Value Judgment Task. This is because the expected response we associate with the existential wide-scope reading (i.e. 1b) will also apply to the distributive reading (i.e. 1c). This restricts considerably our methodological tool box that we can apply to such sentences.

For these reasons, we will divide our review into two parts, the first part involving sentences with an existential quantifier and a universal quantifier and the second part involving sentences with a quantifier and negation.

In section 2, we will look at sentences with an existential and a universal quantifier, like (1) and (2) above. We will find that forced choice questionnaire studies reveal a preference for overt scope in examples like (1), and that a series of self-paced reading tasks showed that inverse scope in examples like (1) comes with a processing cost (Anderson 2004). This can be taken to support a syntactic (or semantic) analysis of such examples where the inverse scope reading is syntactically (or semantically) more complex than the overt scope reading. We will also review a growing body of evidence that found that children do access inverse scope readings in various types of sentences involving scopal ambiguities, even when the same ambiguity is not present in the adult language, i.e. so-called Rigid Scope languages.

In section 3, we will look at examples with a quantifier and negation, like (3) and (4) above. We will demonstrate that such examples are preferentially interpreted with surface scope in a task called the Incremental Verification Task (Conroy 2008). The same preference was found in a speeded force choice task, while the non-speeded force choice task and a sentence completion task yielded no preference for either reading. In addition, we will see that adults have preference for the inverse scope reading with examples like (3) in the truth-value judgment task and that their surface scope bias can be successfully alleviated in the incremental verification task by priming (Conroy 2008). We will consider the so-called Parser Hypothesis, that proposes that adults and children have an intrinsic parsing bias for overt scope. We will also consider the Extra-Linguistic Hypothesis, that assumes that the observed parsing biases are not the result of genuine preferences in the parser for the overt scope reading, but are rather emergent results of the interplay of various extra-linguistic factors, such as how different tasks place different demands on the hearer and how they lead to different verification strategies in different sentences.

Throughout we will consider data from adults and children alongside each other. We will review the literature on the development of scope in first language acquisition. We will review evidence in favour of what has been called the Observation of Isomorphism (Musolino et al. 2000, Lidz & Musolino, 2002), where it was found that in examples like (3) and (4) children have a strong preference for the overt scope reading in comprehension. However, the

significance of this finding has been questioned. First, Gualmini (2004) and Hulsey et al. (2004) found that the inverse scope reading is available to children if the pragmatic conditions are favourable. Second, Conroy et al. (2009) found that the isomorphism effect found is only present in 5-year olds, but not 4 year-olds, giving rise to a U-shaped developmental curve.

Overall, we will conclude in section 4 that the evidence is mixed but ultimately comes down in favour of the Extra-Linguistic Hypothesis. Thus, the hypothesis that is most consistent with the body of evidence as a whole is that children, like adults, have access to both the overt and the inverse scope reading from the start. Various extra-linguistic factors, and potentially task effects, must then be responsible for the fleeting observation of isomorphism in children.

2 Universal-existential combinations

2.1 Adult psycholinguistic evidence

2.1.1 Early work: evidence for Overt Scope Preference

In this section we will investigate sentences involving an existential and a universal quantifier, such as (1) and (2) repeated here for convenience. Recall that the overt scope reading in both is the one where the scopal order of the quantifiers corresponds to their surface syntax c-command relations, i.e. (1b) and (2b). The so-called inverse scope reading is the one where the semantic scopal order is the opposite of the surface syntax c-command order, i.e. (1c) and (2c). Recall also that the reading where a group of individuals are involved for both quantifiers, i.e. (1c) and (2b) is called the distributive scope reading.

(1) a. A doctor advised every nurse.
b. a > every: There is a doctor that advised every nurse.
c. every > a: For every nurse there is a doctor that advised

(2) a. Every nurse assisted a doctor.
b. every > a: Every nurse assisted a doctor, but not necessarily the same doctor.
c. a > every: There is a (specific) doctor and every nurse assisted him/her.

Early work on such sentences argued for the relevance of linear order (Van Lehn (1978), Fodor (1982)): the earlier the quantifier appears in the sentence, the wider the scope it has. Ioup (1975) proposed that the relative position of the two quantifiers on the argument hierarchy determines their scopal relation. Reinhart (1983) proposed that only surface c-command relations matter, which are not easily distinguished from linear order in a right-branching language like English. Yet others, such as Kempson and Cormack (1981) and May (1985) argued for extras-syntactic factors, such as the topical nature of the quantifier to play a role, in the sense that topics take wide scope over nontopical quantifiers. Finally, it has also been noted that different quantifiers exhibit

differing degrees of likelyhood for taking wide scope over other quantifiers (see e.g. Ioup 1975, Kroch 1975). (See Tunstall (1998) for an extensive overview).

Most of these studies involved corpus investigations or relied on paraphrases offered by the participants or on various metalinguistic judgments by the participants. For instance, Catlin and Micham (1975) asked participants to say which noun phrase the sentence was 'about' and the members of which noun category would one need to examine in order to determine whether the sentence was true. Some studies involve complex judgment tasks: for instance, Micham et al. (1980) asked participants to determine whether a given sentence was true in a situation depicted by a diagram or table matching two sets of participants corresponding to the two quantified arguments, not unlike our own sketch in (6) above. All of these methodologies raise important questions. One has inadequate control over the experiment if subjects can offer their own paraphrases. Metalinguistic judgments, as involved in Catlin and Micham's study, and problem-solving tasks, as in Micham et al. (1980), are likely to involve central cognitive processes to a higher degree. Our intention is to determine whether native speakers access inverse scope readings when they use language naturally. So, it is not desirable to get them to calculate the readings as a kind of 'maths' problem. Nevertheless, bearing their methodological baggage in mind, it is important to note that all of these studies found some support for a general preference for overt scope. So, people generally preferred to interpret such sentences with semantic scopal relations corresponding to surface c-command relations of the quantifiers.

This can be formulated as a broad and intuitive hypothesis, which we will simply state here as follows:

(9) Overt Scope Preference (OSP)
When a scopal ambiguity arises, people have a preference for assigning the overt scope interpretation over the inverse scope interpretation.

Kurtzmann and MacDonald (1993) were the first to manage to successfully ease the methodological tension between obtaining convincing results about the nature of the reading assumed by the participant without overburdening them (and potentially interfering with the results) with a metalinguistic task. Rather than asking metalinguistic judgments or presenting participants with a problem-solving tasks, they provided disambiguation sentences with singular or plural subjects (e.g. *the kid*… vs. *the kids*…) which were presented to the participants after they read a doubly quantified sentence such as *A kid climbed every tree*. Their line of thinking was that if participants obtain overt scope in the doubly quantified sentence they will favour a continuation referring to a singular subject referent, while inverse scope would favour a plural subject in the continuation. They also included control items that had either unambiguous surface scope (e.g. *The same kid climbed every tree*) or unambiguous inverse scope (e.g. *A different kid climbed every tree*). Their results indicated the interplay of a variety of factors including syntactic scope between the quantifiers, but also the relevance of thematic roles, verb type (activity vs. stative) and topicality (i.e. subjects are preferred topics and take wide scope).

Tunstall (1998) replicated their study and adapted it to three-argument verbs as in *Kelly showed a photo to every critic last month*. *The photo was*…/ *The photos were*…. This way she successfully eliminated the topicality and verb-type factors. This is because the subject, which is the default topic in these sentences, is not quantificational. The question of scopal order concerns the direct and the indirect objects. The method used was a self-paced word-by-word stops-making-sense reading task, meaning that participants were asked to indicate if the sentence combination they are reading stops making sense to them. She hypothesised that participants will have a preference for overt scope, along the lines of our OSP.[3]

Interestingly, the results diverged. In the *a… every*-condition (i.e. items like (1)), participants had a significantly longer reading time in the critical region (subject and auxiliary) of the continuation sentence if the subject was plural compared to if the subject was singular (Tunstall 1998: 66). So, participants favoured the overt scope reading in these sentences. However, no comparable difference was found in the *every... a*-condition (i.e. in items like (2)). Here, Tunstall did not find longer reading times in the critical region for singular subjects. She proposed to account for this asymmetry by her Vagueness Principle (Tunstall 1998:71):

(10) Vagueness Principle (Tunstall 1998: 71)
> When the processor gives *every* wide scope over an indefinite, it can remain vague (underspecified) as to whether the indefinite is multiply instantiated or not. This information can be filled in by further inferencing or by subsequent context.

So, her interpretation is that all the results are consistent with OSP, i.e. in *every... a* sentences too, *every* takes wide scope over the existential, reflecting their relative c-command relations, but this does not force people to choose a plural continuation. So, you do not see a cost for continuation sentences starting with a singular subject in this case. In other words, Tunstall distinguished overt vs inverse scope from distributivity. The latter is not necessarily a direct consequence of the former: a universal can take wide scope over an indefinite without necessarily triggering a distributive interpretation of the indefinite. Thus, Tunstall could maintain the generality of the OSP despite the diverging results across the *a… every* and *every… a* conditions. Note that an alternative interpretation of the data would be to take the divergence of the results at face value and propose different grammatical analyses for the two quantifier orders. We will come back to this idea below.

2.1.2 The role of context: Evidence from self-paced reading studies: Anderson (2004)

Let us now review a series of experiments that addressed the question whether in online processing the parser has a preference for overt scope in *a...every* or

[3] Tunstall's (1998:56) actual *Principle of Scope Interpretation* is a more technical formulation of the *Overt Scope Preference.*

every...a sentences, and whether such a preference can be mitigated or perhaps even reversed by appropriate discourse context. The tentative conclusion we will draw is that there is a robust overt scope preference for *a...every*-sentences, while the evidence of an overt scope preference for *every...a*-sentences is murky at best. We will also see that context does influence the availability of the inverse scope reading, but that in most cases, again for *a...every*-sentences, the parser's preference for overt scope cannot be totally obliterated. These conclusions will lead us to question Tunstall's uniform analysis of the two types of sentences.

Anderson (2004) performed a series of experiments to investigate parsing preferences for doubly quantified sentences. Her first experiment was a questionnaire study using items like (11). It was a variation on Kurtzmann & MacDonald's (1993) methodology in the sense that it involved two sentences, one with the two quantifiers and a paraphrase that would disambiguate the reading.

(11) *Example test item for questionnaire study* (Anderson 2004:32 ex 47)
A cashier greeted every customer.
a. One cashier greeted customers.
b. Several cashiers greeted customers.

The singular paraphrase was chosen 81% of the cases, indicating a strong preference for the overt scope reading.

In her second experiment, Anderson (2004) decided to investigate whether contextual bias influences people's parsing preferences. So she tested the same sentences as in the first experiment embedded in a context that biases for overt scope (12a) and inverse scope (12b).

(12) *Example test items for questionnaire study with context* (Anderson 2004: 35, exs. 53-54)
a. *Overt-scope-biasing context*:
The members of the gourmet club decided to put out a cookbook of their favorite recipes. They wanted the recipes to be easy enough for an inexperienced cook. The president of the club requested that a volunteer test the recipes to make sure that the instructions were correct. After a short discussion, **a member of the club tested every recipe**.

b. *Inverse-scope-biasing context*:
The members of the gourmet club decided to put out a cookbook of their favorite recipes. They wanted the recipes to be easy enough for an inexperienced cook. Members who nominated recipes were required to test the recipes to make sure that the instructions were correct. **A member of the club tested every recipe**.

Anderson also included two control conditions, which used the same contexts as the ambiguous conditions, but which employed test sentences that unambiguously indicated an overt scope reading or an inverse scope reading. For the former she used a definite NP subject (i.e. *The helpful member tested every recipe*.), for the latter she used a distributive modifier (i.e. *A different*

member tested every recipe.) Each paragraph was followed by a forced choice comprehension question, as in (13).

(13) How many club members tested recipes?
One. Several.

The results revealed an effect of context as the singular response for the comprehension question, indicating an overt scope interpretation, was chosen 81% of the time in the ambiguous overt-scope biassing context but only 47% of the time in the ambiguous inverse-scope biassing context (Anderson 2004: 40). (The corresponding findings for the unambiguous conditions were 99% in the overt-scope context and 4% in the inverse-scope context.) Note that the overt-scope biassing context did not seem to have a positive effect in the sense that comparable response rates were found in Experiment 1 without contextual embedding. Nevertheless, there was still a strong bias towards an overt scope interpretation, despite the presence of a context with an inverse-scope bias.

Anderson also performed the experiment for *every...a* items, but the results are not comparable, as she used a rating scale instead of a forced choice response. The results showed an effect of context in that case too. They also showed that the plural paraphrase was acceptable to people as it had an average rating of 3.3 on a 5-point scale, which is significantly higher than the mid-point at 2.5, while the surface scope paraphrase had a 4.3-rating (Anderson 2004: 43).

Anderson (2004) also performed a self-paced reading task with sentence pairs adopted from Kurtzmann & MacDonald's (2003) study:

(14) *Example test items from self-paced reading study* (Anderson 2004: 45)
a. An experienced climber scaled every cliff.
The climber(s) was/were very skilled.
b. Every historian examined a document.
The document(s) was/were in good condition.

The residual reading times calculated for the continuation sentence as a whole were significantly longer in the *a...every* condition, if the continuation sentence had a plural subject compared to when it had a singular subject (Anderson 2004: 48). (There was also a marginal difference in the same direction on the final part of the continuation sentence, i.e. the part following the finite verb.) This indicates a processing cost associated with inverse scope, because participants found it more taxing to read a continuation sentence that is compatible with an inverse scope reading, suggesting that they had committed themselves to an overt scope interpretation.

Interestingly, no comparable (but opposite) difference was found for the *every...a*-sentences. Here the reading times for the continuation sentence as a whole or at any of its regions were not longer if the subject of the continuation sentence was singular, compared to when it was plural (Anderson 2004: 52). This is compatible with at least two interpretations. First, and this is what Anderson proposes, we can follow Tunstall's line of thinking: There is an overt scope parsing preference in these cases too, but since that does not commit to

a plural subject in the continuation sentence (see Vagueness Principle), the task is inadequate to find this difference. In other words, participants may have chosen the singular response and nevertheless entertain an overt scope interpretation.

But the data is also compatible with an alternative explanation, namely that in this case, there is no overt scope parsing preference in the first place. The second explanation can perhaps be further supported by looking at the answers participants gave to the comprehension questions in (13). As we can see in Table 1, (Anderson 2004: 49, Table 3; 54 Table 5) participants chose the response for the comprehension question that was compatible with an overt scope reading in around 90% of the time in both the *a...every* and the *every...a* conditions when the disambiguating sentence was compatible with that reading (i.e. in the singular subject in the *a...every* condition and the plural subject in the *every...a* condition). But they behaved differently in the face of a continuation sentence compatible with inverse scope. In the *a...every* condition they still chose a response to the comprehension question that indicated an overt scope interpretation 59% of the time, giving evidence for a reluctance to entertain inverse scope. In contrast, in the *every...a*-condition they only did so 18% of the time. To me, this shows that they were not averse to an inverse scope reading in this case, as they chose a response to the comprehension question indicating that reading 82% of the time.

	a... every	***every...a***
Singular subject disambiguating sentence	87%	18%
Plural subject disambiguating sentence	59%	91%

Table 1: Percentage of surface scope response for comprehension question (i.e. *One.* in *a...every* condition; *Several*. in *every...a* condition) for the two types of continuation sentences

Anderson also tested the sentence pairs in (14) in biassing contexts like the ones in (12). She used a 2 x 2 design with quantifier order and contextual scope bias as the two controlled variables. There were two interesting findings. First, in the *a...every* condition the continuation sentence that disambiguated for the inverse scope reading (i.e. the one with a plural subject) was read more slowly than the continuation sentence that disambiguated for the overt scope reading (Anderson 2004: 61-62). So there was a processing cost involved with the inverse scope interpretation for *a...every* sentences even if the context was biased for that interpretation. It is interesting to note that the processing cost could only be measured on the continuation sentence, as the reading times for the critical regions of the doubly quantified sentences did not reveal any relevant differences. There was a marginal effect of context, with the inverse scope-biassing context leading to marginally longer response times irrespective of the type of disambiguating sentence (singular subject or plural subject).

Second, unlike in the experiment without context, marginally longer reading times were found for the continuation sentence that disambiguates for the inverse scope reading (this time that is the one with singular subject). There are

two possible explanations for the difference between the experiment without biassing context and the one with context. Anderson takes it to mean that an inverse-scope biassing context boosted participant's willingness to entertain an inverse scope reading. In turn, the higher proportion of inverse scope interpretations led to an increased processing cost. But in my opinion, it is not clear that this was in fact the case. This line of thinking would predict that there would be a discrepancy in the proportion of inverse scope responses in the *every...a*-condition depending on whether the context was biassing for overt scope or inverse scope. However, this was not the case: the proportion of responses indicating an inverse scope interpretation (i.e. *One*. in the *every...a*-condition) was 71% in the overt-scope biassing context and 69% in the inverse-scope biasing context (Anderson 2004: 66 Table 7). So, there is in fact no evidence that contextual bias had an effect on participants' willingness to entertain inverse scope for *every...a*-sentences. Given that the result was only marginal, it seems more parsimonious to conclude that there is no conclusive evidence that there is a parsing preference for the overt scope interpretation in *every...a*-sentences.

The final pair of experiments that Anderson conducted addressed the question whether the increased processing cost measured on the continuation sentences indicating inverse scope in the *a...every*-condition was the result of the parser committing to an overt scope representation early on and the continuation sentence forces a reanalysis, which leads to processing cost, or whether participants in fact do not commit to a resolution of the scope ambiguity until the continuation sentence, and then the measured processing cost can be attributed to deriving the inverse scope reading. In one experiment (Anderson 2004: 69), she tested *a...every*-sentences with no discourse context and no continuation sentence in a self-paced reading task. The sentences were always followed by a comprehension question, see (13) above. She analysed the reading times data separately depending on whether participants' response to the comprehension question indicated an overt scope reading or an inverse scope reading. Calculating residual reading times for the entire sentence, a significant difference was found: participants who entertained the inverse scope reading read the sentences significantly slower compared to participants whose response to the comprehension question revealed an overt scope interpretation. Anderson (2004: 73) concludes, that it is the assignment of the inverse scope reading to the sentence and not a subsequent reanalysis that presents a processing load.

In her final experiment, Anderson (2004: 73-74) used the same *a...every*-sentences, this time embedded in biasing contexts. Again, there were no continuation sentences and the data was divided in two based on the response participants gave to a follow-up comprehension question. She also included two unambiguous conditions, using items like in her second questionnaire study described above. The unambiguous sentences were only embedded in matching discourse contexts. The results revealed that reading times were significantly longer in the inverse-scope biassing context compared to the overt-scope biasing context. This was true for ambiguous as well as unambiguous items. She concludes that the favourable context did not mitigate the processing cost associated with entertaining the inverse scope reading. This

goes against the principle of parsimony (Altmann & Steedman 1988, Crain & Steedman 1985), which states that in case of parsing ambiguities, a reading that fits with the discourse context better is always favoured. In contrast, it seems that a processing cost was incurred even in the presence of favourable context. Moreover, the processing cost was not even alleviated when the item itself received unambiguously inverse scope (i.e. items with the modifier *different*), so essentially, even if there is no actual ambiguity between the inverse-scope reading and the overt-scope reading.

Overall, regarding the tentatively hypothesised *Overt Scope Preference*, we can conclude that the reviewed psycholinguistic evidence presents a robust overt scope preference for adult processing of *a...every*-sentences, but not for *every...a*-sentences. It turns out that it is possible to mitigate this preference using a discourse context that biases the reader towards an inverse scope interpretation, but the preference does not fully go away. Finally, we have seen that the real-time assignment of inverse scope comes with a processing cost in *a...every*-sentences, but the same was not so clearly present in *a...every*-sentences.

2.1.3 Theoretical considerations regarding the adult psycholinguistic evidence

The question I would like to consider in this subsection is what the theoretical implications are of the above adult psycholinguistic findings. It is the nature of theoretical discussions that they inevitably engage with technical details of sometimes complex proposals. Some readers of this chapter will relish the thought, others less so. I can reassure the latter type of reader that they can fully profit from this chapter by omitting to read this subsection and jumping straight to section 2.2 to consider the review of the data from language development.

Let us briefly consider what kind of theoretical proposals are available in the literature to account for scopal ambiguities. In generative syntax it is generally assumed, following May (1977), that inverse scope in *a… every*-sentences is obtained by a covert movement operation, quantifier raising, henceforth QR, as illustrated in (16) for a sentence like (7).

(16) $[_{IP}$ every building $[_{IP}$ an American flag was $[_{VP}$ hanging in front of $t_{NP}]]$

It was shown that QR, like other instances of A-bar movement, is island-sensitive and it may give rise to semantically distinct readings.[4] QR, as its name suggest, can only effect generalised quantifiers. Existentials like the object noun phrase in examples like (2) do not have that option. Rather, they obtain wide scope by different means. Different proposals exist with respect to the nature of existential wide scope. For ease of exposition, let us adopt Reinhart's (1997) choice function account for existentials. The details of the choice

[4] In fact, QR seems to be even more strongly local than A-bar movement, as it is generally clause-bound. The interested reader is referred to Reinhart (1997) for a historical overview of different conceptualisations of QR.

function mechanism need not concern us here. What matters for us here, is that these assume that existentials do not take wide scope via a syntactic movement operation.[5] Reinhart (1997) showed that such a distinction (i.e. QR vs. choice function) is empirically justified given the divergent grammatical properties of wide scope indefinites and wide scope universals: Wide scope indefinites, for instance, are not island-sensitive, and generally have properties that liken them to *wh*-in-situ. Wide scope universals are island-sensitive, even clause-bound, and in general their properties are similar to those of moved *wh*-elements.

Other theories do not posit any kind of asymmetry between how the grammar assigns overt versus inverse scope. In Steedman's (2000) Combinatory Categorial Grammar, for instance, the different scopal possibilities are derived by differential orders of composition between the verb and the noun phrases in question. If the verb combines with the object first and then the subject, the subject will take higher scope. If the verb combines with the subject first, and then the object, the object will take scope over the subject. The former composition gives rise to what we call overt scope and the latter to the inverse scope reading. There is no asymmetry at the level of syntactic structure, in fact in this system the terms 'overt' and 'inverse' make little sense. There is also no asymmetry of type, in the sense that indefinites and universals are treated in the same way by the system.

Tunstall's own (1998) proposal sits halfway between the two. She posits an asymmetry in terms of how syntax and semantics map onto each other, but does not distinguish between the different types of quantifiers. In this respect, Tunstall's proposal is similar to the theoretical proposal put forward by Bobaljik & Wurmbrand (2011). They proposed that there is an interface principle in the grammar that ensures that surface syntax c-command relations are respected at LF, and consequently in the semantics, by corresponding scopal relations. These two proposals are essentially more technical formulations of our *Overt Scope Preference* idea.

So, we have three different approaches to quantifier scope interactions: one that treats overt scope and inverse scope on a par (e.g. Steedman's framework), one that treats overt scope as preferable in general (e.g. Tunstalls' and Bobaljik and Wurmbrand's proposal), and one that treats overt scope preferable in those cases where inverse scope would necessitate QR. Other formulations of these theoretical possibilities are available in the literature, but for ease of exposition I singled out these proposals to illustrate these three logical possibilities.

Considering the predictions of these theories for psycholinguistic data, note that the three proposals are asymmetrically entailed. Steedman predicts no asymmetries, Bobaljik & Wurmbrand predict a general overt scope preference,

[5] Note that the same applies arguably numerals also as in (4). Numerals can be generalised quantifiers, in which case they can obtain wide scope via QR, or they can be existentials, in which case they take wide scope by other means.

while Reinhart predicts a preference for overt scope in those cases where inverse scope would be the result of QR. In other words, this means that Bobaljik & Wurmbrand's theory would not be falsified by data that would support Reinhart's proposal, it would simply have to be enriched to provide an account for the unexpected difference between the parsing preferences associated with the different type of quantifiers. Similarly, if it turns out that there is a general parsing preference for surface scope, that would not be incompatible with Steedman's framework. Rather, some additional mechanism would need to be invoked to account for the observed difference. The predictions have a stronger bite in the opposite direction. Reinhart's (1997) proposal would be called into question if a general rather than a particular overt scope preference was found to be present in parsing; and Bobaljik & Wurmbrand's proposal (as well as Reinhart's) would be questioned by findings that point to a systematic lack of evidence for any kind of parsing preference for overt scope.

In this light, we can review the findings enumerated in section 2.1.2. Anderson's findings are by and large compatible with Reinhart's (1997) predictions. There was a clear overt scope preference in *a…every*-sentences, which are purported to involve QR by Reinhart (1997), but not in *every…a*-sentences, which Reinhart assumes do not involve QR. Nevertheless, there were also more general tendencies for overt scope present in all the conditions, and the fact that this overt scope preference could only be partially mitigated by context is also important to note. This, if taken at face value, would provide support Tunstall's and Bobaljik & Wurmbrand's approach.

Anderson's (2004) findings also revealed that in *a…every*-sentences, but not so much in *every…a*-sentences, the parser experiences an extra processing load when considering inverse scope readings, even if that reading is supported by context and moreover, even if the utterance is unambiguous. One question that arises is why the construction of an inverse scope interpretation would present a processing load. It is not the case that instances of A-bar movement generally have this effect. Reinhart (1997, 2006) argued that the processing load is due to the fact that QR involves global economy considerations, which are costly to the processor. This might go some way to explain the findings, but under this view one would potentially expect the processing cost to be diminished if the inverse scope reading is supported by the discourse context, and perhaps even eliminated altogether when the sentence in question takes an inverse scope reading unambiguously. The robustness of the processing cost in such sentences supports a general default overt scope interpretation, which the parser seemingly has to abandon if faced with the inverse scope reading. In a way that is not dissimilar to the reassignment of the syntactic position of the PP-modifier in sentences like *I saw the man with the binoculars* from a VP-adjoined position to a NP-adjoined position in the course of the parse. Whether this has to do with the presence of an existential in the subject position, or the default topical nature of the existential subject, or a genuine overt scope preference, should be explored further in future research.

Let us now turn to the findings from language development. As we will see, these findings also have interesting theoretical implications, which we will also discuss.

2.2 Evidence from child language acquisition

2.2.1 A surprising result: scopal freedom

Based on the adult findings reviewed above, we can expect that children would also have an overt scope preference for sentences involving an existential subject and a universal object. In fact, given that Anderson found that the inverse scope reading incurs an extra processing load in *a…every*-sentences, we might even expect children to have an exaggerated preference for overt scope in these sentences. Whether this is in fact so depends in large part on what actually causes the processing load associated with inverse scope in such utterances. We will come back to this issue at the end of the section, where we will consider the theoretical implications of the language development findings.

The literature on quantifier scope interactions rarely considers existential-universal quantifier pairs. This is due the presence of an entailment between the two readings, which, as we have already noted above makes it methodologically difficult to test such sentences using tasks that rely on the assignment of a truth value. Nevertheless, a pioneering study on quantifier scope interactions involved precisely these quantifiers.

Japanese is a so-called Rigid Scope language where inverse scope by quantifier raising is severely restricted. A sentence like (17a), for instance, would be assigned overt scope by adult native speakers. In this language, the distributive scope reading (every > some), would be available in utterances with scrambling, like (17b). Here the object c-commands the subject in surface syntax, so the distributive reading can be obtained without resorting to covert scope.

(17) a. Dareka-ga daremo-o sementa
someone-NOM everyone-ACC criticized
“someone criticized everyone” (unambiguous)
b. Daremo-o_i dareka-ga t_i semeta
everyone-ACC someone-NOM criticized
Lit. “Everyone, someone criticized” (ambiguous) (Goro 2007: 57-58 ex.41)

Goro & Akiba (2004), reported in Goro (2007), performed a truth-value judgment task with English and Japanese children and adults using sentences with an existential subject and a universal object, as in (18).

(18) Dareka-ga dono tabemono mo tabeta (Goro 2007, 47-48, ex. 36)
someone-NOM every food ate
'Someone ate every food.'

The story involved an eating contest with twelve group of animals. Each group consisted of three animals of the same type (e.g. 3 pigs, etc.). Each group was

invited to eat three different pieces of food (e.g. a cream puff, a banana and a pepper). The child was told that there are two important rules of the game. First rule is all the food must be eaten. Second rule is each of the group members has to eat something. So, if each member of a group eats exactly one item, the group wins and gets a gold medal. If one animal is greedy and eats up all the three food items, the group gets a black cross, i.e. symbol of failure. Also, if they all refuse to eat one of the food items, they also get a black cross. The outcome of the story was that 4 groups performed according to the rules and received gold medals. 4 groups had a greedy member who ate up all the food and thus received a black cross and 4 groups shared the food out nicely but the final member didn't finish eating their food item, so they ended up receiving a black gross too.

The critical trials are the groups that received a gold medal as here the inverse scope reading of the test sentence is true while the overt scope reading is false. In the second batch with the groups with greedy animals, the reverse is true. While the third batch involves groups that attempted to perform what corresponds to the inverse scope but failed to do so.

16 Japanese children with a mean age of 5;4 (range: 4;10-5;9) accepted the inverse scope reading, so answered YES to the test item 42.2% of the time. A group of 16 Japanese adults never accepted the critical test items. 16 English-speaking children with a mean age of 5;4 (range 5-0-5;10) accepted the inverse scope reading 35.9% of the time, while 29 English-speaking adults did so 33.6% of the time (Goro 2007:53, ex. 38). Goro (2007) also performed a control study with different test items, to ascertain that Japanese children also have access to the overt scope reading of such sentences.

Goro (2007) concluded that Japanese adults revealed an unwillingness to assign inverse scope to such sentences. This matches previous findings in both the theoretical and the empirical literature on Japanese about Japanese being a Rigid Scope language. At the same time, Japanese children's behaviour patterned with that of English-speaking children, and *not* with that of Japanese adults.

Szendrői et al. 2017 found very similar results with German children and adults using an actout task. One advantage of the actout task is that it does not require truth-value judgments on the part of the participant, thus making it ideal for testing existential-universal quantifier pairs. At the same time, it is important to note that methodologically speaking, the actout task is less than ideal to investigate any kind of ambiguity. This is because the participant displays their preferred reading in this task. It is possible, therefore, that the actout task would underrepresent all the readings that the participant would be able to assign to the test sentence.

German is also a Rigid Scope language, with scrambling. There too, adults showed scope rigidity with utterances like (19), assigning an overt scope reading 98% of the time. In contrast, 20 5-year-old children with a mean age of 5;3 (range: 5;1-5;7) performed an inverse scope action 56% of the time and 20 6-year olds (mean: 6;4, range: 6;1- 6;11) did so 42% of the time.

(19) Ein Tierpfleger füttert JEDE Giraffe. (Szendrői et al. 2017:xx ex. 13)
A zookeeper feeds EVERY giraffe.

This shows that the unexpected scopal freedom that Goro found with Japanese children is not a language specific effect, but rather, the same holds in another Rigid Scope language too. (See also Zhou & Crain 2009 on Mandarin discussed below.)[6]

2.3 Interim summary

Recall that in the adult psycholinguistics literature, it was found that overt scope is preferred in *a...every*-sentences, unlike *every…a*-sentences, and that inverse scope in the former type of sentences comes with a processing cost. The evidence from child language paints a different picture. Here we saw that even in languages with Rigid Scope, 5- and 6-year old children showed scopal freedom. We cannot offer a satisfactory reconciliation of these facts at this point.

Instead, let us turn to scope interactions involving negation. This should be useful to investigate whether the overt scope preference of the parser is something that holds more generally, or whether perhaps it arises as an emergent phenomenon due to the nature of the doubly-quantified sentences we explored so far.

[6] In terms of a more detailed theoretical perspective, it is possible to think of Scope Rigidity in the form of cross-derivational, global economy, as suggested by Reinhart (1997). In this line of thought, the reason why German and Japanese SVO sentences lack the inverse scope reading would be precisely that these languages allow for alternative word orders (i.e. e.g. scrambling) that have the same distributive scope reading without recourse to inverse scope. These alternative orders effectively block the availability of inverse scope reading (See also Bobaljik & Wumrbrand (2011) for the same point). If children fail to carry out such global cross-derivational comparisons due to processing limitations, as Reinhart (1999, 2004) suggested, then they would be expected to fail to exclude the inverse scope reading of the SVO utterances. They cannot retrieve the alternative word order variant that obtains the distributive reading under overt scope and compare them under the intended interpretation, so they have no reason to exclude the inverse scope reading, so no blocking takes place. Hence their lack of scope rigidity. This is how Szendroi et al proposed to account for their results. Note, however, that Goro (2007) argues against such a blocking account and offers an alternative account based on the conversational implicature of maximality associated with the Japenese particle *ga*. Persuasive as his account is, however, it would not easily carry over to German. So, we shall have to leave the issue of why children in rigid scope languages consistently experience scopal freedom open for future research.

3 Scope interactions of quantifiers and negation

3.1 Setting the scene

Recall that Tunstall (1998) proposed that the parser is intrinsically endowed with a preference for overt scope. Anderson's (2004) findings endorsed this, at least for *a...every*-sentences. But is it in fact an intrinsic property of the human parser to preferentially assign overt scope? Or is it perhaps the case that the parser considers both overt and inverse scope and extra-linguistic factors influence the final choice resulting in an outcome that prefers overt scope? So, this preference is an emergent consequence of the combination of extra-linguistic factors and the properties of the parses, not due to an intrinsic property of the parser itself.

One possible way to explore the generality of the parser's overt scope preference is to conduct a variety of tasks with the same sentences. If the parser has a general preference, this should show up in all or most tasks, or at least in the tasks that tap into early preference. We will review experimental evidence that has been amassed on various tasks both off-line and online ones, with or without context. The findings put together show a mixed picture, with some tasks showing a strong overt scope preference, while others not.

Another issue that could shed light on this issue is the comparison of data from adults and children. Assume that the parser has an intrinsic preference for overt scope, which can be overriden in favourable contexts in the case of adults, although there is some evidence that even in this case the inverse scope reading incurs a processing cost. In this scenario it would be a natural extension of the state of affairs in adults that children would have an even stronger preference for surface scope. It is well-established that children have smaller working memory resources and also that they are generally less able to capitalise on at least certain types of discourse-contextual information (e.g. Noveck 2001). Both would point in the direction that children's ability to override the assumed overt scope preference of the parser should be diminished compared to adults' ability, resulting in an even more robust overt scope preference. In contrast, it is also possible that the parser has no intrinsic preference, but rather supplies both overt and inverse scope readings. It could be the result of a combination of grammatical and extra-grammatical factors that ultimately adults show an overt scope preference in many tasks, especially those without supporting discourse context. In such a scenario, whether children show an overt scope preference would depend on their knowledge of the relevant grammatical factors and their susceptibility to the relevant extra-grammatical factors. If in both domains that are adultlike, then we would expect children to have either the same behaviour as adults, so we would expect them to show an overt scope preference in many tasks. Alternatively, if children either lack necessary grammatical knowledge or are less susceptible to the relevant extra-linguistic factors, then we would in fact expect a less robust preference for overt scope compared to adults.

In the following sections we will review a number of studies. The reader might find it helpful to refer Table 2 for details of each experiment.

Table 2: Summary of experimental findings reviewed in this paper with the exception of Anderson's (2004) studies

Test item	Source	Surface syntax	OS	IS	% YES Adults	%YES 4yo	% YES 5-6yo
The detective didn't find some guys	Musolino et al. 2000	neg > exist	NO	YES	100	35 (4;7)	65 (5;7)
The Troll didn't deliver some pizzas.	Gualmini 2004	neg > exist	NO	YES		90 (4;10)	
Every horse didn't jump over the fence	Musolino et al. 2000	univ > neg	NO	YES	100		7.5 (5;11)
	Musolino & Lidz 2006	univ > neg	NO	YES	92.5		15 (5;4)
Every bug didn't hide behind the tree.	Viau et al. 2010	univ > neg	NO	YES		22.3/ 38.8 (4;5), 80.6 primed	
Every bunny didn't eat a purple carrot.	Conroy 2009 Q/A task	univ > neg	NO	YES	8/20 100; 9/20 0		
Every dwarf didn't spraypaint the barn that belongs to the pig/ the cow	Conroy 2009 sentence completion task	univ > neg	pig	cow	40?= 10/20 100 'pig', 10/20 86.6 'cow'		
Every dwarf didn't spraypaint the barn that belongs to the pig/ the cow	Conroy 2009 speeded FC	univ > neg	pig	cow	18.5 'cow'		
Every dog isn't wearing a hat	Conroy 2009 IVT	univ > neg	last	1st	22.9 1st (incl. 14/22 0)		
Every cow doesn't have a hat	Conroy 2009 IVT	univ > neg	1st	2nd	63.4 1st (with 10/20 100)		
Every cat didn't hide behind the sofa.	Conroy et al. 2009	univ > neg	NO	YES	76	81 (4;9)	44 (5;4)= (7/15 0, 8/15 82.5)
Every horse jumped over the log, but every horse didn't jump over the fence	Musolino & Lidz 2006	univ > neg	NO	YES	100		60 (5;4) = 6/10 100, 4/10 0
Every horse jumped over the log, but every horse didn't jump over the fence (Mandarin)	Zhou & Crain 2009	univ > neg	NO	YES	0	89 (3;4-4;3)	10 (4;5-5;11)
			YES	NO	100	100 (3;4-4;3)	100 (4;5-5;11)
The detective didn't find two guys	Lidz & Musolino 2002	neg > num	NO	YES	93	33 (4;4)	
The Troll didn't deliver two pizzas	Gualmini et al 2008	neg > num	NO	YES		75 (4;6)	
The detective didn't find two guys	Musolino & Lidz 2003	neg > num	YES	YES	75 OS, 7.5 IS, 17.5		
John didn't find two hearts	Conroy 2009 IVT	neg > num	early	late	unclear		
John didn't find two hearts	Conroy 2009 IVT	neg > num	late	early	47 early; 40 late		
Two frogs didn't jump over the rock	Musolino & Lidz 2003	num > neg	NO	YES	27.5		

Ein Tierpfleger füttert JEDE Giraffe	Szendrői et al. 2017 actout	exist > univ			2 IS		56 IS (5;3), 42 IS (6;4)
Dareka-ga dono tabemono mo tabeta	Goro & Akiba 2004,	exist > univ	NO	YES	0		42.2 (5;4)
Some ate every food	TVJT	exist > univ	NO	YES	33.6		35.9 (5;4)

3.2 The 'Observation of Isomorphism': evidence from Truth-Value Judgment Tasks

Musolino et al. (2000) tested 15 children with an average age of 4;7 (range 3;10-5;2), 15 children with an average age of 5;7 (range: 5;2-6;6) and a group of adults in a truth-value judgment task, using sentences like (14).

(20) The detective didn't find some guys.

The context story involved a situation where different characters hid behind various objects and the detective's task was to find them. The outcome of the story was designed to satisfy the inverse scope reading (i.e. There was someone the detective didn't find.) but falsify the overt scope reading (i.e. There was at least one guy the detective found.) Not also, that the overt scope reading is independently ruled out because '*some N*' is a positive polarity item in English, so it must take scope over negation. The adults accepted the test sentence 100% of the time, while the older children did so in 65% of the time, and the younger ones 35% of the time (Musolino et al. 2000: 10). All these results were significantly different from each other. Children's justification for their NO response was that the detective did find someone, so they revealed an overt scope interpretation, despite *some N* being a positive polarity item in adult grammar.

Musolino et al (2000) also tested 20 children with an average age of 5;11 (range 4;0-7;3) and a control group of adults on sentences like (21). In the story three horses attempt to jump over a barn, but they realise it is too high for them to jump over, then they decide to jump over a fence. Two horses jump over the fence, but the third one fails to do so. This outcome makes the inverse scope reading (not> every) true, while the overt scope reading (every > not) is false.

(21) Every horse didn't jump over the fence.

Children accepted the test sentence in 7.5% of the time, while adults did so 100% of the time. This indicated a very strong preference for overt scope with these types of sentences too.

There is an asymmetric entailment relation between the two different scopal possibilities of a negation and a universal in the sense that the reading where the universal takes wide scope over negation entails the reading where the negation takes wide scope over the universal. In other words if it is true that none of the horses jumped over the fence (i.e. every> not), then it is also true that not every horse did so (i.e. not > every). For this reason it is impossible to make the every >not reading true in a situation while making the not> every reading false. This poses a methodological problem for the truth-value judgment task, as this task relies on associating the reading which is by assumption harder to obtain with the YES answer, and the other reading with the NO answer. This is not a problem in utterances like (21) where the targeted inverse scope reading is the one where negation takes wide scope over the universal. But in sentences like (22), one cannot associate the every> not

reading with YES, while making sure that the not > every reading will correspond to a NO response.

(22) The detective didn't find every guy.

Lidz and Musolino (2002) got around this problem by testing sentences like in (23). Given that numerals do have a quantificational meaning (alongside an existential one, which we can put to the side here) such sentences can test whether the utterance in (23) can be interpreted distributively with the numeral taking scope over negation. But this time it is possible to create a situational context where the overt scope reading (not > two) is false. For instance if the detective tries two find four guys, and manages to find two of them but not the other two, then it will be true that there are two guys such that the detective didn't find them (i.e. two > neg), but it is false that he found less than two, as he did in fact find two people (i.e. neg > two).

(23) The detective didn't find two guys.

Like before, Lidz and Musolino (2002) found that 24 English-speaking children with a mean age of 4;4 (range 3;11- 4;11) accepted such sentences in the given context 34% of the time, while a group of 24 adults did so 93% of the time (Lidz & Musolino 2002: 131-132). Again this shows an overt scope preference by children, albeit a milder one than before, while adults are able to access the inverse scope reading.

To sum up, in a series of truth-value judgment tasks involving sentences with various quantifiers and negation it has been found that children have a preference for the overt scope interpretation. This has been termed the 'Observation of Isomorphism' (Musolino et al. 2000: 14). Adults, in contrast, were able to access the inverse scope reading in all these cases.

3.3 Possible research hypotheses

Let us investigate this effect further. There are essentially four possible reasons for its existence. First, it is possible that children have a grammatical deficit. They simply have not yet acquired the grammatical tools that underlie inverse scope (e.g. quantifier raising). Second, it is possible that children and adult have the same grammatical knowledge, but children's parsers unlike adults' has an intrinsic preference for overt scope. This would mean that children's and adults' parsers are qualitatively different, and thus we would need to find the so-called magic moment, when children mysteriously abandon their child parser and turn into adults. This approach goes against the spirit of the Continuity Hypothesis (Pinker 1984, Crain & Thornton 1998) and thus should only be considered if the other approaches fail to account for the data. Third, it is possible that both children and adults are capable of deriving inverse scope grammatically speaking, but their parser has an intrinsic preference for overt scope. Adults' parsing preference for overt scope is exaggerated in children due to their limited memory resources. This would mean that there is no qualitative difference between adults and children. Both populations have an intrinsic parsing

preference for overt scope, but this is more pronounced in children. Following Conroy's (2008) work, let us call this the Parser Hypothesis.[7] Fourth, it is possible that children and adults are capable of deriving inverse scope and they do not have an intrinsic parsing preference for overt scope. Rather, extra-linguistic factors are responsible for the semblance of an overt scope preference. Such factors could have differing effects in different experimental tasks and also interact in interesting ways in different age groups. So, it is possible that children's appearance of an overt scope preference plays out differently from adults' appearance of an overt scope preference. Conroy (2008) termed this the Extra-Linguistic Hypothesis. Let us review these hypotheses in turn.

3.4 The Grammatical Deficit hypothesis

It is easy to see that the first option can be dismissed. Let us now demonstrate that. Lidz & Syrett (2004) tested 24 4-year-olds (range 4;1-4;10) in a between subject's design on sentences that involve and ambiguous VP-ellipsis site, such as (24a). Such sentences involve Antecedent Contained Deletion. The quantificational object *every X* involves a VP-ellipsis cite. VP-ellipsis is normally resolved under the Parallelism Constraint: the elided VP is the same as its antecedent. But in sentences where the VP-ellipsis is inside the object that is inside the antecedent VP this leads to infinite regress. The solution is to assume that the quantificational object *every X* undergoes QR to the position where it c-commands the material that has been elided (Fiengo & May 1994, Merchant 2000 and others). In a sentence with embedded clauses like (24), its position thus determines the size of the elided VP: if the QR adjoins to the embedded verb, as in (18b), the antecedent of the elided VP will be interpreted as the embedded VP, and if adjoins to the matrix verb, as in (24c), then the elided VP will be interpreted as the matrix VP.

(24) a. Miss Piggy$_i$ wanted to PRO$_i$ drive every car that Kermit did.
b. Miss Piggy wanted to [$_{vP}$ [$_{DP}$ every car that Kermit did <*drove t*>]$_i$ [$_{VP}$ drove t_i]]
c. Miss Piggy [$_{vP}$ [$_{DP}$ every car that Kermit did <*wanted to drive t*>]$_i$ [$_{VP}$ wanted to drive t_i]]

Lidz & Syrett (2004) found that if the context story was consistent with an embedded reading and falsified a matrix reading, children gave a NO response indicating a matrix reading 54% of the time while adults did so 32% of the time. In most cases both children's and adults' justification of their NO responses revealed a genuine matrix reading. So, we can conclude that children possess the grammatical knowledge to apply quantifier raising as a syntactic operation, just like adults.

[7] Conroy (2008) in fact named this hypothesis the Parsing Hypothesis, but both me and an anonymous reviewer finds that name less intuitive given what it means, so I changed the name for the hypothesis to be Parser Hypothesis.

A further reason to doubt that the Observation of Isomorphism is due to a deficient grammar is that children appear to be able to access the inverse scope reading in other experimental setups. Gualmini (2004) was the first to notice that information structuring can alleviate children's reluctance to assign inverse scope in sentences involving negation and an indefinite or numeral object. They argued that children could indeed access inverse scope if the reading with inverse scope provided an appropriate answer to what they called the 'question under discussion'. In particular when the expectation is built up that the Troll should deliver all the pizzas, and he ends up delivering two, but loses two, children were no longer unable to access the inverse scope reading of *The Troll didn't deliver some/two pizzas*. In the experiment with *some*, children's inverse scope responses jumped from 50% in Musolino (1998) to 90% in Gualmini's (2004) experiment. In the experiment with *two*, children's inverse scope responses jumped from 50% in Musolino (1998) and 33% in Lidz & Musolino (2002), to 75% in Gualmini's (2004) experiment.

Gualmini *et al*. (2008) argued that this substantial improvement occurred because the expectations of the situation make the question 'Will the Troll deliver all the pizzas?' highly accessible, and the inverse scope reading (i.e. Two/Some pizzas were not delivered.) is a more appropriate answer to this question than the overt scope reading (i.e. The Troll didn't deliver any/(at least) two pizzas.) So, children do appear to consider inverse scope when the information-structure requirements of the story require them to do so. In particular, Gualmini et al.'s view is that children access whichever reading provides a felicitous answer to the so-called Question Under Discussion, which is an abstract construct that maintains information flow in discourse. This points towards a scenario where children have no grammatical deficit. Gualmini et al.'s specific explanation is most consistent with the idea that neither children nor adults have an intrinsic parser preference for overt scope either. Rather, extra-linguistic factors sometimes cause adults and more frequently children to favour the overt scope reading in some experimental tasks, i.e. Conroy's 'extra-linguistic hypothesis'.

Musolino & Lidz (2006) also provided evidence against a lack of grammatical knowledge in children, when they demonstrated that children who fail to access the inverse scope reading in examples like (21), repeated here for convenience, nevertheless do so in examples like (19).

(21) Every horse didn't jump over the fence.

(25) Every horse jumped over the log, but every horse didn't jump over the fence.

They tested 20 English-speaking children (8 boys and 12 girls) between the ages of 5;0 and 5;11 (mean 5;4) and 20 adults on sentences like (21) and (25) in similar contexts favouring an inverse scope reading. They found a significant difference for acceptance rates for children, 15% for utterances like (21) and 60% for utterances like (25). (Note that this was a bimodal distribution of 6/10 children accepting the test sentence 100% of the time and 4 children rejecting it 100% of the time.) Adults, in contrast, accepted the test sentence 92.5% and

100% of the time, respectively. Their explanation for children's improvement is reminiscent of Gualmini's explanation: it is children's immature pragmatic abilities that stop them from displaying their correct grammatical knowledge in certain scenarios. Once the pragmatic conditions are favourable, as in (25), the children are able to access the inverse scope reading. In fact, as Viau et al. (2010) demonstrated, it is not the actual contrast in the test sentence that makes the inverse scope reading shine through, but rather the difference in the events enumerated in the context stories. In the stories that tested (25) the horses first successfully jump over an obstacle (i.e. the log) before they attempt the jump over the fence that only some of them manage. It is the presence of this early success in the story that proved to be the relevant factor, and not its explicit mention in the test sentences.[8]

One final argument against a grammatical deficit account of the Observation of Isomorphism comes from a priming study. Viau et al. (2006) tested 4-year-old children in a priming task with utterances like (26a) and (26b). One group of children received 6 instances of an utterance like (26a), while the second group of children received three such utterances preceded by three instances of utterances of the type illustrated in (26b).

(26) a. Every bug didn't hide behind the tree.
b. Not every bug hid behind the tree.

For the first group, proportion of inverse scope judgments were 22.25% for the first three utterances and 38.8% for the last three utterances. In contrast, the proportion of inverse scope judgments for the first three utterances in the second group was 83.3% and 80.58% for the last three utterances. Viau et al. (2010) interpreted their findings to show that utterances with an unambiguous distributive scope facilitate the distributive inverse scope interpretation in ambiguous utterances. This is a sort of semantic priming effect where certain aspects of the meaning of an utterance prime the same aspect in an utterance which can optionally have that interpretation. Assuming that children are not able to attain readings that are beyond their grammatical competence, we may assume that priming successfully nudged their processor to consider the inverse scope reading that they seemed to be initially unable to do so.[9]

---

[8] Interestingly, the children tested by Viau et al. (2010) were 4-year-olds. This age group acted differently in Conroy et al.'s (2009) experiment where they had 81% acceptance in stories with no early success event. We will come back to this point below.

[9] The same effect was demonstrated for adults by Conroy (2008). She performed a series of experiments which established that priming occurs for the unexpected interpretation in both children and adults (inverse scope for children and surface scope for adults) but adults' priming effects can be modulated according to immediately previous exposure. Although these results are interesting in their own right, as they reveal interesting aspects of semantic priming, ultimately, our lack of understanding of the underlying mechanisms of semantic priming make it difficult to draw any firm conclusions with respect to scope interpretations.

In fact, the priming effect worked in a rather more subtle way too. Viau et al. (2010) also tested utterances with a universal subject and negation in stories with early success and no early success. First, they reconfirmed that if children heard the test sentences with three stories with early success, their performance was boosted compared to those children that heard the same stories without early success (50% vs. 25% inverse scope readings). But then all the children heard three stories with no early success. In these stories, the proportion of inverse scope reading was 80% for the children who had heard early success stories before, and remained 25% for those that heard no early success stories. This means that simple exposure to a story with a discourse setup that favours inverse scope boosted performance not only for the children that actually showed sensitivity to this discourse manipulation in the first three stories, but also for some of those children that did not reveal a sensitivity to the discourse manipulation earlier. This suggests that for some children the sensitivity was there, but they were not fast enough to integrate the discourse information to provide a matching judgment.

Overall, based on the above studies we can conclude that children's grammar is not deficient. They are able to perform the syntactic operation of quantifier raising and they can even use it to obtain inverse scope readings, they are just reluctant to do so in some experimental tasks. But their performance can be boosted by various pragmatic manipulations and by semantic priming.

3.5 The Parser Hypothesis

3.5.1 Evidence from TVJT-tasks

Musolino & Lidz (2003) considers the position, discussed third above, that both children and adults have an intrinsic preference for overt scope. This is what I termed the Parser Hypothesis, (see Conroy 2008). Recall that Lidz and Musolino (2002) found that adults accessed the inverse scope reading 93% of the time for sentences like (23) repeated here for convenience when the sentence was presented in a context that was compatible with the inverse scope reading and falsified the overt scope reading.

(23) The detective didn't find two guys.

In Musolino & Lidz (2003), they tested the same sentences in contexts that were compatible with either reading. Their findings revealed that adults' justification indicated an overt scope interpretation 75% of the time. This shows that although adults, unlike children, can access the inverse scope interpretation of such sentences, they nevertheless have a preference for the overt scope interpretation.

Similarly, in sentences with a numeral subject and negation, like (27), adults no longer showed the overwhelming ability to access the inverse scope reading that they demonstrated with sentences involving a universal subject (i.e. 100% acceptance for items like (21) above from Musolino et al. 2000). Using sentences like (27) in a context that favours the inverse scope reading and

falsifies the overt scope reading, they found that adult participants only accepted such sentences in 27.5% of the time, indicating a substantial reluctance to access the inverse scope reading (Musolino & Lidz 2003:9).

(27) Two frogs didn't jump over the rock.

Musolino & Lidz (2003) concluded that one way to explain these facts is if adults too have an intrinsic preference for overt scope in such sentences, albeit they are able to revise their initial parse to fit a context that favours the inverse scope reading, while children lack the processing resources to do so. But there are also a couple of lose ends. First, adults' performance on sentences with a numeral subject and negation like (27) differed from their performance on sentences like (23) with a numeral object and negation, while children's performance was uniform on both. Lidz & Musolino (2003) propose that this may be because the sentences have different underlying grammatical mechanisms: A sentences like (23) involves quantifier raising of the object over negation, while (27) is more likely to involve reconstruction of the subject to a position under negation.[10] Potentially, this explains the difference for adults, although given the underlying assumption that children have the same grammatical capacity and parser preferences as adults, it is not clear how it follows that children's performance is bad on both.

At the same time, it is also important to note that adults perform differently on sentences involving negation and a universal subject, like (21), compared to sentences with negation and a numeral subject, like (27). If both involve reconstruction, why do adults find the sentences with a numeral subject more difficult. Lidz & Musolino (2003) give a potential explanation. They note that one crucial difference between such sentences is the entailment relations that hold between the overt and inverse scope readings in the universal case but not the numeral one. They argued (see also Lidz & Musolino 2006) that for sentences with universal subjects the every>not reading is more efficiently expressed by an utterance such as *None of the horses/ no horse jumped over the fence.*, so the hearer can reason that if the speaker used *Not every horse jumped over the fence* then they are more likely to have meant the every> not reading, because otherwise they would have used the more efficient and unambiguous utterance *No horse.../None of the horses....* In this way Lidz & Musolino (2003) attribute adults' preference for the inverse scope reading to Gricean reasoning, while underlyingly they have access to both readings. Children, they hypotesise, may have limited processing resources that stop them from engaging with such Gricean reasoning (see also Reinhart 1995, 2006 for the same claim), hence the lack of preference for the not> every scope in their case.

But there is one issue raised by this account. If adults show an inverse scope preference in sentences with universal subjects and negation as a result of Gricean reasoning, then the question arises, why the same reasoning cannot be invoked to explain a potentially diametrically opposite preference. There is,

[10] But compare with Reinhart (2004) who argues that neither sentences have QR.

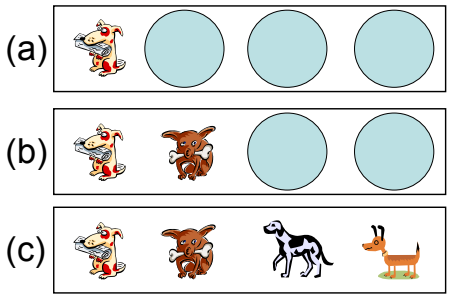
of course, also a more efficient and unambiguous way to express the not>every-reading, namely by using an utterance such as *Not every horse jumped ....* It is not clear why adults are sensitive to one potential alternative but not the other.

3.5.2 Evidence from IVT-tasks

It is equally possible that this inverse scope preference for adults is in fact not generalisable, and to some extent it is a special consequence of the truth-value judgment task. To explore this possibility, let us turn to a novel task, the Incremental Verification Task, designed by Conroy (2008). One crucial difference between the IVT and the TVJT used in the experiments described above is that the IVT does not involve a full-fledged discourse context.

The IVT task invites participants to judge if a sentence is true in a picture as soon as they feel they have enough information to judge. The picture itself has four subparts which are each hidden under a cup. Participants can reveal a growing proportion of the picture by removing cups from left to right one after the other. An example item with a picture is given in (28).

(28) Every dog isn't wearing a hat.

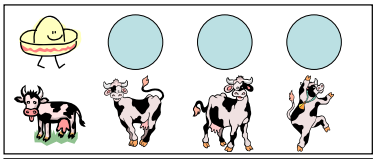


In this item the inverse scope reading can be verified after removing the first cup (28a), while all the cups must be removed for the verification of the overt scope reading (28c). Both readings are true in the picture, but one can still distinguish which reading a participant entertained by checking how many cups they removed to reach a decision. Participants entertained the overt scope reading (i.e. persisted to the last cup) in 77.1% of the time, with 14/22 participants doing so 100% of the time (Conroy 2008: 56).

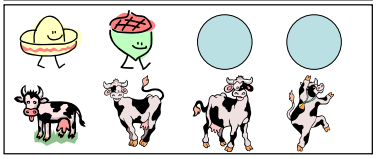
One might wonder whether perhaps participants favour the overt scope reading for some task-specific reason, like for instance that they persist with removing cups until the reading that requires the largest amount of information can be verified. But this seems unlikely given that Conroy also tested items like (29). In these items, the cows but not their possessions are all visible from the start. In such items, the overt scope reading can be verified (or more precisely, falsified) after the first cup is removed (see 29a), while two cups must be removed to verify the inverse scope reading (29c).

(29) Every cow doesn't have a hat.

(a)



(b)



In such trials, participants entertained the overt scope reading 63.4% of the time, with 10/20 participants entertaining it all the time. But we know that adults adhere to the inverse scope reading in truth-value judgment tasks where the inverse scope reading is associated with the YES answer (e.g. 96.6% (Conroy 2008: 121); Mussolino et al. 2000 reported above around example (21); also Lidz & Musolino 2006). There are at least two important factors to consider as to why this difference between the TVJT and the IVT occurs. First, the high proportion of YES responses in the TVJT is likely to be boosted by what Crain & Thornton (1998: 212) call the *Principle of Charity*, which states that participants always respond YES in a truth-value judgment task when the reading associated with the YES answer is available to them. The same issue does not arise in the IVT where there is no compelling reason that would cause participants to settle on a judgment early or late in the task. Recall that for sentences with negation and universal objects, the truth-value judgment task that associated the inverse scope reading with YES and the overt scope reading with NO yielded 93% inverse scope responses (Lidz & Musolino 2002), while the same task yielded 75% overt scope responses once both readings were associated with a YES-response (Musolino & Lidz 2003). So, the effect attributable to the *Principle of Charity* is rather large indeed. [11]

[11] Note also that one important difference between the IVT and the TVJT is the the former does not include a discourse context, while the latter does. In fact, in this light, Conroy's (2008) IVT findings can also be interpreted as supporting Lidz & Musolino's (2003, 2006) account based on Gricean reasoning. This is because it would make sense for the Gricean reasoning to apply in a fully-fledged discourse but not in what one could describe as a situation of uncertainty, such as in the IVT. Chierchia et al (1998) and Gualmini et al (2003) argued that tasks that require a verificational judgment before the whole discourse situation is known force participants to take decisions in 'prediction mode', which has the effect that scalar implicatures will be cancelled. For instance if someone told you that 'There will be pizza or ice cream at the party' before the party takes place and in the end there was both pizza and ice cream at the party, you would not say they pronounced an untrue statement. But if they were to utter the sentence 'There was pizza or ice cream at the party' *after* the party had taken place, then you would think they are giving an imprecise account of what happened (Crain et al. 2000). This is because cancellation of the scalar implicature that provides the exclusive reading for *or* is justified in a situation of uncertainty, i.e. before the party, but not in a situation of full

To sum up, given Conroy's findings using the IVT, it seems that there is in fact an early parsing bias for the *overt* scope in sentences with universal subjects and negation (77% vs 63% in the two experiments). This would be in line with Tunstall's (1998) proposal that the parser has an intrinsic overt scope preference, and our own OSP. But is this effect reproduceable in other tasks, or is it perhaps some emergent effect that is the result of task effects associated with the IVT and other extralinguistic factors?

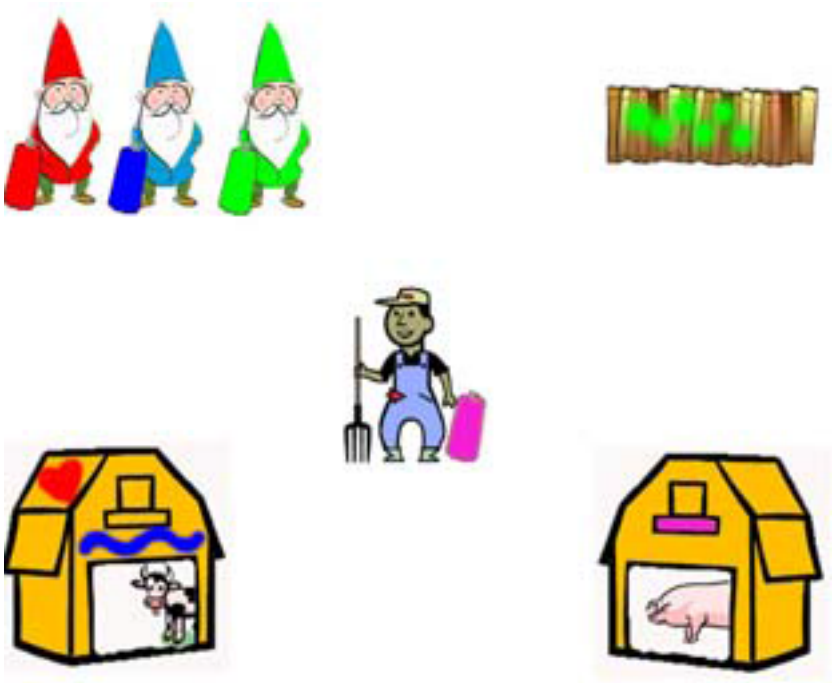
3.5.3 Evidence from Forced Choice tasks

Conroy performed two further tasks to probe this question further. She performed a sentence completion task and a speeded force choice task using the same stimulus discourse contexts that allow for a felicitous use of both scope readings, see (30). The sentence fragment for the sentence completion task is given in (31). The only difference in the speeded force choice task was that participants were instructed to choose one of the two pictures of the pig's or the cow's barns and they were instructed to do that as soon as they can. While there was no time pressure in the sentence completion task.

(30) *Example context and image for sentence completion task and speeded force choice task*

Here, there is a red, blue and green dwarf, with their cans of spraypaint. The farmer has pink spraypaint. There is a barn that the cow lives in, and a barn that the pig lives in. It looks like the red and blue dwarves spraypainted the cow's barn, but not the green dwarf. It doesn't look like any of the dwarves spraypainted the pigs barn, so the farmer finished the job.

knowledge, i.e. after the party. Arguably, the IVT consists of such a situation of uncertainty, as the participant's task is to verify a sentence in an unfolding situation (i.e. during the decision process more and more aspects of the situation are revealed). So, I speculate that this task would cancel scalar implicatures of the type that Lidz & Musolino (20003, 2006) hypothesised to boost inverse scope readings in the TVJT.



(31) Every dwarf didn’t spraypaint the barn that belongs to the ...

If participants entertain an overt scope reading of the universal subject and the negation, they will opt for the pig's barn as none of the dwarves painted that. If they entertain an inverse scope reading, they will choose the cow's barn, as all but one of the dwarves painted that. (The stimuli were counterbalanced for physical positioning and temporal mention order effects.) In the non-speeded sentence completion task 40% of the responses indicated an inverse scope reading (Conroy 2008: 93). But this was based on a bimodal distribution of half the participants never accessing the inverse scope reading and the other half accessing it 86.6% of the time (Conroy 2008: 93). We note that this should result in an overall average of 43.3% inverse scope, not 40%, so there must be a typo in the original text somewhere. In contrast, a significantly different result was obtained in the speeded force-choice task where participant's overall rate of inverse scope choice was 18.5% (Conroy 2008: 93). Conroy's interpretation is that the results of the sentence completion task are in line with her previous findings in the IVT task: there is overall a mild preference for overt scope. But note that the preference was much less pronounced, in fact it could be as low as 56.7%. Let us also note that there was a bimodal distribution of some participants consistently going for overt scope and some consistently entertaining inverse scope. This weakens the conclusion that there is in fact an intrinsic parsing preference for such sentences. At the same time, the speeded task revealed that under pressure, participants are overwhelmingly more likely to settle for the overt scope reading, which constitutes a fairly strong argument in favour of an overt scope parsing preference.[12]

[12] Conroy also performed the IVT task using numerals and negation (Conroy: 2008, 58-63), using sentences like *John didn't find two hearts*, participants entertained the overt scope reading in 47% of the time in those trials where the

Conroy (2009) also performed a judgment task with unbiassed context like in (32). Participants' task was to answer a question, as in (33), indicating their scope judgment. This task was slightly different in terms of its discourse information structure, as here, unlike in the sentence completion task, the actual outcome of the story was not revealed in the context. The task was also more similar to a TVJT task in that a YES/No answer was required from participants.

(32) There was a party at Farmer Jon's farm. A bunny from Hillsdale, a bunny from Stonybrook and a bunny from Camelot came. Farmer Jon offered carrots all around, but they were purple. He also had some cauliflower. Although the bunnies were all hungry, each one thought that purple carrots might not taste too good and considered eating the cauliflower instead. But, there was a lot more of the purple carrots and Farmer Jon kept saying how good they were. He really hoped that they would all try them. But in the end, **every bunny didn't eat a purple carrot**. At the end of the day, the bunnies had a cool glass of celery juice to drink.

(33) Did some bunnies eat a carrot?

There were 12 target paragraphs. The results showed a strongly bimodal distribution with 8 out of 20 participants never obtaining an inverse scope interpretation, and 9 doing so all of the time. The obtained results were thus very similar to the non-speeded sentence completion task described above.

3.5.4 Taking stock

Conroy puts forward two different hypotheses to explain her data. Under the Parser Hypothesis, which posits an intrinsic parser preference for surface scope reasoning, one could posit that adults would have a parsing preference

---

overt scope reading could be verified earlier, and they entertained the overt scope reading in 40% of the trials where the inverse scope reading could be verified first. The difference between the two types of trials were not significant, and there were 8/22 participants across trials that only entertained the overt scope reading. Conroy concluded that participants only have an overt scope preference in this task with sentences involving universal subjects and negation but not with numeral objects and negation.

Let us compare this with results reviewed above using the truth-value judgment task. Musolino & Lidz (2003) found an overt scope preference (75%) in a task associating both readings with a YES answer using similar sentences. They do not report individual data. But given Conroy's report of a bimodal distribution and given that all these studies involve a relatively low number of participants (normally 20), it seems reasonable to conclude that a clear overt scope bias has not been demonstrated overall for items involving negation and a numeral object. At the same time, one issue that certainly leaves room for thought is the consistently bimodal distribution of scope judgments found in these experiments.

for the surface scope at an early stage in the comprehension process, which is reflected in their results in the ITV task. This overt scope preference is later revised to match the situational and discourse context. This would explain the adult preference for the inverse scope in the TVJT task. In a speeded task there is time pressure on the participant, which could arguably stop them from revising their interpretation, hence the predominantly overt scope response in that condition. Conroy herself notes that this hypothesis does not explain the difference between the bimodal distribution of responses found in the non-speeded sentence completion task, and the unimodal inverse scope pattern found in the TVJT. Both tasks are non-speeded and involve a full-fledged discourse context, although the sentence completion task makes both readings true, while the TVJT only makes the inverse scope reading true. There is, of course, one more difference between the two, namely that the Principle of Charity biases towards in inverse scope interpretation in the TVJT but not in the sentence completion task.

In contrast, the Extra-Linguistic Hypothesis would posit that both the surface and the inverse scope readings are available to the parser, which does not have an intrinsic preference for either. Conroy claims that this easily explains the results of the sentence completion task. It does indeed do so, to the extent that both readings are manifested.[13] Under the extra-linguistic hypothesis, any results that show an overt scope bias or an inverse scope bias need further explanation. We have already provided one for the TVJT results with sentences involving universal subjects and negation. As Lidz & Musolino (2003) explain, the two scopal readings of such sentences are in an asymmetric entailment relation so in a TVJT we can expect that adults (but not children) perform Gricean reasoning, favouring the inverse scope reading, even though they do not have an intrinsic parsing bias for it. But how to explain the overt scope bias found in sentences with universal subjects and negation with the IVT?

3.5.5 Two strong arguments against the Parser Hypothesis

Macdonald et al. (1992) showed that adults with low word span recall have difficulty comprehending syntactically complex ambiguous sentences. Carpenter et al (1994) showed that concurrent load can be used to tax working memory resources even in people with normal or high span. Specifically, Waters et al. (1987) found that articulatory suppression affected adults' ability to process syntactically complex sentences. Listening to irrelevant speech impacts participants' word span, indicating that it taxes the phonological loop (Colle & Welsh 1976). For this reason, Conroy (2008: 211) performed a reading task using non-biased context under a concurrent task taxing working memory resources. Their hypothesis was that if the Parser Hypothesis is on the right track, the concurrent load should influence parsing of scopally ambiguous sentences in non-biassed contexts, leading to a higher proportion of overt scope readings. This is because by hypothesis, overt scope is accessed first

[13] Note that I am not sure why this hypothesis would give rise to such a markedly bimodal distribution. The only reason I can think of, is self-priming, but the effect seems too strong for that.

and inverse scope can be obtained as a result of subsequent revision which places a burden on the working memory resources of the parser. In contrast, if it turns out that concurrent load makes no difference to the proportion of overt scope readings obtained, that would question the validity of the Parsing Hypothesis. Her 20 adult participants obtained an overt scope reading 53% of the time, which was not statistically significantly different from the rate of overt scope readings obtained in the baseline condition, without the concurrent task (Conroy 2008: 215). Thus, the findings revealed no effect of concurrent task, even though they pre-tested the task to show that it does indeed impact word span (Conroy 2008: 211). Conroy concluded that the results call the Parsing Hypothesis in question.

A final argument against the Parser Hypothesis is put forward by Conroy et al. (2009). They noted that the children who were susceptible to the discourse manipulation in Gualmini's (2004) study were about a year younger than the children tested in Musolino & Lidz (2006). (There were also other differences. For instance that Gualmini tested negation and numeral objects, while Musolino & Lidz tested universal subjects and negation.) They tested 15 4.5-year olds (4;5-5;2, mean 4;9) and 15 5-year-olds (5;3-5;7, mean 5;4) and 12 adults in a truth-value judgment task using test items like (34) where the outcome of the story is depicted in Figure 1. As shown in the figure, the inverse scope reading was true in the story, while the overt scope reading was false.

(34) Every cat didn't hide behind the sofa.

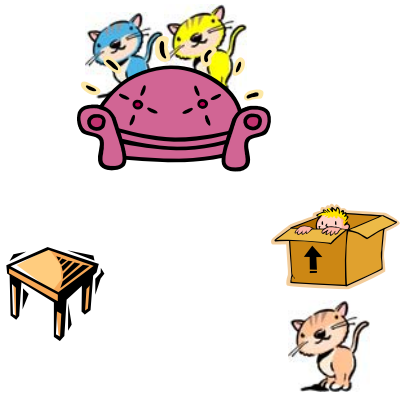


Figure 1 Outcome of example test story from Conroy et al.'s (2009) TVJT task

Adults accepted the inverse scope reading 76% of the time, which is lower than in other similar TVJT tasks. The authors do not have an explanation for this unexpected finding. 4.5-year-olds accepted the inverse scope reading 81% of the time. In contrast, 5-year-olds accepted the inverse scope reading 44% of the time, which was marginally significantly different from the rate of inverse cope readings obtained by adults and significantly different from that of 4-year-olds. The distribution of the 5-year-old participants' data was bimodal, with 7 out of 15 children never accessing the inverse scope reading. (It follows, that the remaining 8 children accepted the inverse scope reading 82.5% of the time.)

Conroy et al. (2009) offer an account for what they term the 'fleeting ismorphism effect' of 5-year-olds, as follows. 'First, because younger children can obtain the inverse scope interpretation, and presumably younger children do not have more parsing resources than older children, we conclude that the isomorphism effect in five year olds cannot be due to immaturity of the sentence parser, as claimed in Lidz and Musolino (2002).' (Conroy et al. 2009: 13). They also 'conclude that the isomorphism effect does not solely derive from a failure to experimentally meet felicity conditions' (Conroy et al. 2009: 13), contrary to Gualmini et al's (2008) conclusions. Rather, they propose 'children adhere to a U-shaped development in the domain of scope ambiguity resolution '(Conroy et al. 2009: 13)

The idea, which is provided in more detail in Conroy's (2008) work, is that at the early stages (i.e. until Age 4.5) children have non-adultlike parsers with an inverse scope preference. The reason is claimed to be that children aim to 'mimick the inverse scope interpretations observed in the input' (Conroy 2008: 146). Later on, at the Age of 5, they acquire an adultlike parser, but at this point, they are still assumed to be 'lacking the ability to revise their interpretation according to discourse information' (Conroy 2008: 146), resulting in an inability to revise their initial interpretation. Only when they are able to appropriately integrate discourse information will they reach the end of the U-shaped curve, and behave in an adult-like way.

A couple of things to note with respect to this explanation are the following. First, the evidence that adults entertain an overwhelmingly higher proportion of inverse scope readings with sentences involving a universal subject and negation rests on a small data set informally collected by Musolino. There are no formal corpus studies based on large spoken or written corpora to back up this assumption. Second, Conroy (2008: 201) herself argues that many examples collected 'in the wild' actually cannot be properly classified, as the discourse context in which they appeared did not disambiguate the two scope readings convincingly. One wonders then how children are supposed to perform this task to arrive at a parser strategy that mimicks adult proportion of inverse scope readings.

Nevertheless, Zhou & Crain's (2009) findings are important to mention here. They tested Mandarin equivalents of sentences with universal subjects and negation in an early success context TVJT. Mandarin is scopally rigid like Japanese. Zhou & Crain's finding patterned very similarly to Conroy et al.'s (2009) findings. They found that the children whose age range was 3;4-4;3 accepted an inverse scope reading 89% of the time, while the older children, who were aged 4;5-5;11 did so only 10% of the time, with adults never accepting the inverse scope reading. In the Mandarin data there is, of course no U-shaped pattern, given that adults disallow inverse scope in such sentences. But it is interesting to note that the drop in inverse scope readings for Mandarin children seems to occur at the same age, around the end of the 4th year of life. Could it be that (some) English children are briefly experimenting with a Rigid Scope parameter during Age 5?

So, overall, there seems to be evidence for 5-year-olds to display a fleeting isomorphism effect in truth-value judgment tasks, displaying what looks like a U-shaped developmental curve. Future research should establish, whether this effect is general in the sense that it is replicable using other tasks, and in the sense that other doubly-quantified constructions also display the effect.[14]

More generally the evidence although mixed, ultimately comes down against the Parser Hypothesis. First, we have seen that an inverse scope bias was found in many adult TVJTs. Comparing these to forced choice tasks and IVTs, however, strongly suggests that the overacceptance of inverse scope is a task effect, due to the Principle of Charity. The results from the forced choice and question-answer tasks revealed no adult preference for overt scope. This argues against the Parser Hypothesis, as does the fact that the so-called Isomorphism Effect in children turns out to be a fleeting one. Perhaps the strongest argument against the Parser Hypothesis is the lack of sensitivity scopal assignments showed for increased working memory load. If we are to abandon the Parser Hypothesis, then the findings that will need to be accounted for are the overt scope preference found in IVT tasks, the overwhelming overt scope preference in the speeded forced choice task and last but not least, the Isomorphism Effect found in children. However fleeting it is, it needs an explanation. In the next section we will review some extra-linguistic factors discussed in the literature that may be helpful in this endeavour.

[14] I would like to note two aspects of the Conroy et al. (2009) task that in my view would merit further investigations. First, unlike in many previous tasks which had 4 test stories, here 6 test stories were performed. This could have partly boosted the effect due to the fact that self-priming has been demonstrated to play a role with scopal judgments (see Viau et al. 2010 reviewed above). In addition, two warm-up stories were administered and 2 filler stories, all of which had the same event structure as the target stories, namely that three characters first failed to perform a particular task, then two out of three proceeded to succeed in a different task, while the third character failed to do so. This is important for two reasons. First, in contrast to Viau et al's (2010) and Lidz & Musolino's (2006) experiments, child participants did not have the discourse advantage of Early Success in this experiment. Nevertheless even 5-year-olds strongly outperformed participants from those studies in comparable No Early Success stories (i.e. 15% inverse scope rate for 5-year-olds in Lidz & Musolino (2006); 22.5% in Viau et al. 2010). The uniformity of the stories is also relevant from the perspective of verificational strategies. In all of the stories an existential verification strategy is a fruitful one (i.e. Find a guy who did/didn't X). This could well be the reason why performance was boosted compared to the other studies. The inverse scope reading requires an existential verification srategy. But training the children to perform existential rather then universal verification would have allowed the possibility that some children choose to apply an existential falsification strategy to the overt scope reading. Of course, this does not in any way provide an answer to the to the intriguing question why about half of 5-year-olds would have decided to do so while no adult or 4-year-old did.

3.6 The Extra-Linguistic Hypothesis

Given the lack of overwhelming evidence for an intrinsic parsing preference for utterances with a quantifier and negation, let us explore why the appearance of such a preference might nevertheless show up in certain tasks.

3.6.1 Verification and falsification

Conroy offers a possible explanation for the overt scope bias found in the IVT task with sentences involving a universal subject and a negation. Assuming that there is no real parsing advantage, she proposes that the overt scope reading is favoured in such examples because of the way participants perform the process of verification for the two readings. She explains that in an utterance involving a universal, say 'Every snail has antennae' people could choose to verify if the sentence is true, i.e. check every snail if it has antennae, or try falsifying it, i.e. check if there is a snail without antennae. The first option involves a universal verification procedure (i.e. checking every snail), while the second involves an existential one (i.e. find one snail such that...). Conroy goes on to argue that both falsification and existential procedures are harder than verification and universal procedures for the human parser.

Now take a sentence with a universal subject and a negation such as *Every snail doesn't have antennae*. In such a situation, participants have a choice to two readings. In the overt scope reading, a verification procedure is available which is also a universal one: one needs to check every snail is without antennae. If so, the sentence is true. The inverse scope reading, however, does not have a verification procedure that is a universal one. One can either use an existential verification procedure, i.e. check if there is a snail without antennae, or one can use a universal falsification procedure, i.e. check every snail if it has antennae. If they all do, then the sentence is falsified. Given that participants did not persist to the last cup in the IVT task when their Truth/False response indicated an inverse scope reading, Conroy concludes that people prefer verification procedures even if they are existential, compared to a falsification procedure, even if that is a universal. But this means, Conroy argues, that perhaps the reason why adults opt for the overt scope reading in IVT is that an easy universal verification procedure is available for this reading, while an existential one must be used for the inverse scope reading.[15]

[15] Conroy (2009) reasons that the verification account would predict that sentences with an existential subject and a universal object such as Anderson's (2004) 'A climber scaled every cliff' should give rise to an inverse scope preference, contrary to fact, as in this case it is the inverse scope reading that has a universal verification process (i.e. check every cliff if a climber scaled it), while the overt scope reading needs an existential verification process (i.e. check to see if there is a climber such that they scaled every cliff). Conroy proposes that the overt scope preference found by Anderson (2004) was in fact due to the topicality of the subject interfering by licensing wide scope for the existential subject. But this fails to explain Tunstall's (1998) findings where

3.6.2 The Semantic Subset Principle

Another extra-linguistic factor that is discussed extensively in Goro's (2007) Crain's (2012) work concerns the effect of the language acquisition device, or more precisely the Semantic Subset Principle on scopal readings in children. These researchers studied sentences with a downward entailing operator, such as negation and disjunction. Due to what is called De Morgan's Law, in logic, a negated disjunction is equal to the conjunction of the negated conjuncts:

(35) $\neg (A \vee B) = \neg A \& \neg B$

As (28) illustrates, this is in fact true in adult English as well. (36) is true in any situation where John brought neither beer nor wine and false otherwise.

(36) John didn't bring beer or wine to the party.

But, interestingly, the same type of utterances in Japanese or Mandarin, has different truth conditions. (37) and (38), the latter a direct translation of (36) are true on the 'not both' reading.

(37) John-wa supeingo ka furansugo-o hanasa-nai
John-TOP Spanish or French-ACC speak-NEG
'John doesn't speak Spanish OR he doesn't speak French.' (Goro 2007: 188, ex. 222)

(38) (Wo cai) Yuehan meiyou dai pijiu huozhe hongjiu qu jiuhui .
(I guess) John not bring beer or wine go party.
'It's either beer or wine that John did not bring to the party' (Crain 2012: 149, ex. 107)

Goro (2007) and Crain (2012) show that this is not the result of De Morgan's Law not holding in the language or the logical connectives having different truth conditions. Rather the readings arise because in Japanese and Mandarin the disjunction takes scope over the negation, so De Morgan's Law does not apply. Thus the LF for the utterances in (37) and (39) is as in (39a and b), respectively.

(39) a. [supeingo ka furansugo-o]$_i$ John-wa t_i hanasa-nai
b. [pijiu huozhe hongjiu]$_i$ Yuehan meiyou dai t_i qu jiuhui.

In Goro & Akiba's (2004) study they tested 30 3- to 6-year-old children (mean: 5;3) using the prediction mode of the truth-value judgment task. In the story there were twelve animals. Each animal was asked, in turn, if it was happy to eat two vegetables, a carrot and a green pepper. The child participants were asked to give the animals rewards as follows: if an animal ate both vegetables

---

topicality as a factor was eliminated. However, I am not convinced in the first place that Anderson's reading tasks actually present a verification problem in the first place. It is possible that the test item is not actually verified (i.e. matched to context to determine its truth or falsity) in the relevant sense. So, no verificational advantage for the overt scope is actually relevant here.

they were supposed to receive a gold medal, if they one of the two vegetables they got a blue medal. If the animal refused to eat both vegetables, they received a black cross, which is a symbol in Japanese culture for failure that that the children were familiar with. After the rewards were given out the puppet uttered the test sentences as a guess. An example is given in (40).

(40) The pig didn't eat the pepper or the carrot.

The critical trials were those where the animal in question had a blue medal. In such trials, as expected, adult controls accepted the test sentence 100% of the time, while children rejected it 75% of the time. In fact, four children were adultlike, and once their responses were removed the rejection rate jumped to 87% for the remaining 26 children. Children's justification revealed an overt scope interpretation, as they interpreted (32) to mean that the pig ate neither pepper nor carrot. Thus, their judgments matched that of English adults (and children) and not that of Japanese adults. A similar study was reported by Crain (2012) with Mandarin children and adults, with the same findings.

Goro (2007) and Crain (2012) argue that the reason children have an overt scope reading in sentences involving a negation and a disjunction in the object is not a general bias towards an overt scope interpretation. Rather, they propose that the Language Acquisition Device helps them avoid a learnability problem. The specific problem is that the two possible readings are in an asymmetric entailment relation. The situations where the 'neither' reading is true are a proper subset of the set of situations where the 'not both' reading is true. As a result, if a child were to initially assume the 'not both' reading, they would run into a learnability problem, given the lack of negative evidence in child language acquisition. Since adults around them might have a grammar that assigns the 'neither' reading (e.g. English) to these operators, such children would never receive positive evidence that would lead them to revise their over-permissive grammar. In contrast, if LAD ensures that children always start out with a the subset grammar, the one with the stronger reading (i.e. the 'neither' reading), they will eventually run into positive evidence that would push them to revise their grammar if they happen to acquire a language where adults use such sentences in the 'not both' sense, such as Mandarin and Japanese.

So, Goro (2007) and Crain (2012) demonstrated that in particular cases of linguistic ambiguity where one reading asymmetrically entails the other, children are expected to entertain the reading that is true in a smaller set of possible situations, i.e. the subset grammar, so as to avoid having to rely on negative evidence to revise their grammar. This is yet another case of an extra-linguistic factor, specifically a factor associated with the language acquisition device, guiding children's interpretation of scopally ambiguous utterances.[16]

[16] Interestingly, as we already mentioned above in a different context, Crain & Hamburger (1992) argued that adults will often have the opposite 'strategy'. In situations where the discourse context is too poor to guide the resolution of a particular ambiguity, adults often adopt the weaker reading, the one which is true in a larger set of circumstances. This is a cooperative discourse move, as in this case, the speaker who uttered the original ambiguous utterance is

4 Conclusions

Overall, let us try to make some helpful broad-brush generalisations based on such a wealth of data involving both existential-universal combinations and interactions of quantifiers with negation, in both adults and children. The evidence is complex, but it seems to me that scopal freedom is default unless quantifier raising is involved, which is in a very restricted set of cases. The appearance of an overt scope preference in the other cases (i.e. with sentences involving negation and a quantifier), I would like to suggest, is more likely an illusion. Scopal readings are evidently very sensitive to different task effects: We have witnessed a high acceptance rate for inverse scope in the truth-value judgement task, at least for adults, but turns out also for most children, except 5-year-olds. We have also seen that the Incremental Verification Task gives rise to an overt scope preference. Although, we can only conclude that for adults for the moment, as the IVT has not been performed with children yet. In addition, scopal readings seem to be easily influenced by priming and even self-priming too, making it even harder to pin down any intrinsic preferences for one scopal reading or the other. But forced choice tasks and unbiassed context question-answer tasks revealed scopal freedom, at least for adults.

On a theoretical level, we may note, that there was not necessarily any theoretical reason to expect an overt scope preference in sentences involving negation and a quantifier in the first place. Reinhart (1997, 2006) argues that such sentences only ever involve quantifier raising if a universal quantifier is c-commanded by negation (i.e. in object position). All the other cases, including all the test items reviewed in this section involve optional reconstruction of the subject under negation, or indefinites or numerals taking wide scope over negation. Reinhart argued that the mechanism for wide scope for indefinites is different from quantifier raising, as it is not island sensitive. If this is all on the right track, then we can return to the data reviewed in Section 2 involving existentials and universals, which, if they occur in this order in surface syntax, and only then, involve quantifier raising. It seems then, that the psycholinguistic evidence supports the assumption of a distinct mechanism for such sentences, i.e. quantifier raising, as in the case of such sentences we found evidence of extra processing load and an overt scope preference for adults. Although, note that the same preference was not present for children in the limited evidence available.

---

assumed to be committed to a weaker statement. In the course of the subsequent discourse, the hearer will have a chance to clarify if the stronger reading was in fact intended. This strategy has precisely the opposite outcome compared to that of the subset principle in children. One example for this effect is for instance the default focal interpretation out of context for utterances with *only* in a sentence where *only* occupies a VP-adjoined position such as 'Peter only gave a book to Sue.' Here adults have a default interpretation where the *only* associates with the indirect object (i.e. Peter didn't give a book to anyone else). 4-5-yer-old children, in contrast, prefer a VP-focus interpretation (i.e. Peter didn't do anything else.) See Szendrői (2004) for details.

In this chapter, I hope to have provided readers with a helpful overview of a very interesting and growing area of psycholinguistics and language acquisition, interpretative ambiguities arising from scopal interactions. I hope to have demonstrated that this is an area where psycholinguistic evidence and evidence from language development can be directly relevant for theoretical analyses of the phenomena. I have also discussed that this area can posit serious methodological challenges. I hope that the readers of this chapter feel well-motivated and better equipped to tackle these research questions in future work. Future research should target sentences with potential quantifier raising more specifically, to find out if the interesting contrast found (adults: processing cost, overt scope preference; children: no preference) holds more generally, or it is specific to sentences with existential subjects and universal objects. Cross-linguistic studies should also be encouraged as they help distinguish cognitive and grammatical factors. Finally, experiments with methodological clarity and ones that test a wide age range promise to be useful to further our understanding, but thenagain they always are.

References

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. Cognition, 30, 191-238.

Anderson, Catherine. 2004. *The Structure and Real-Time Comprehension of Quantifier Scope Ambiguity.* Evanston: Northwestern University dissertation.

Bobaljik, Jonathan D. & Susi Wurmbrand. 2012. Word Order and Scope: Transparent Interfaces and the 3/4 Signature. *Linguistic Inquiry* 43(3). 371–421.

Boster, C. T. and S. Crain (1993) On children's understanding of every and or. Conference Proceedings: Early Cognition and the Transition to Language. Austin, TX: University of Texas at Austin.

Catlin, J., & Micham, D. L. (1975). Semantic Representations as Procedures for Verification. Journal of Psycholinguistic Research, 4(3), 209-225.

Cinque, G. (1993) “A Null Theory of Phrase and Compound Stress,” Linguistic Inquiry 24, 239–298.

Chierchia, Gennaro, Stephen Crain, Maria Teresa Guasti, and Rosalind Thornton 1998. “'Some' and 'or': A study on the emergence of Logical Form.” In Proceedings of the 22nd Boston University Conference on Language Development, 97-108. Somerville, MA: Cascadilla Press.

Chierchia, G., S. Crain, M. T. Guasti, A. Gualmini and L. Meroni (2001) The Acquisition of Disjunction: Evidence for a Grammatical View of Scalar Implicatures, Proceedings of the 25th Boston University Conference on Language Development, 157-168, Somerville, MA: Cascadilla Press.

Conroy, Anastasia. 2008. The role of verification strategies in semantic ambiguity resolution in children and adults. College Park, MD: University of Maryland dissertation.

Conroy, Anastasia, Eri Takahashi, Jeffrey Lidz & Colin Phillips. 2009. Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry* 40(3). 446–486.

Crain, S. (2012). The Emergence of Meaning. Cambridge: Cambridge University Press.

Crain, S. and H. Hamburger (1992) "Semantic Knowledge and NP Modification," in R. Levine, ed., Formal Grammar: Theory and Interpretation, Volume 2, University of British Columbia Press, Vancouver, British Columbia, Canada.

Crain, S. and C. McKee. 1985. The acquisition of structural restrictions on anaphora. In Proceedings of NELS 16. Amherst, MA: GLSA, University of Massachusetts, 94-111.

Crain, S., W. Ni and L. Conway (1994) Learning, parsing, and modularity. In C. Cliifton, Jr., L. Frazier and K. Rayner (eds.), Prospectives on Sentence Processing, 443-466. Amsterdam: Lawrence Erlbaum.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: the use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen & A. M. Zwicky (Eds.), Natural Language Parsing: Psychological, Computational and Theoretical Perspectives (pp. 320-358). Cambridge: Cambridge University Press.

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

Fiengo, R. and R. May. 1994. Indices and Identity. Cambridge, Mass.: MIT Press.

Fodor, J. D. (1982). The mental representation of quantifiers. In S. Peters & E. Saarinen (Eds.), Processes, Beliefs and Questions (pp. 129-164). Dordrecht: Reidel.

Gennari, S., A. Gualmini, S. Crain, L. Meroni, and S. Maciukaite (2001) "How Adults and Children Manage Stress in Ambiguous Contexts," Proceedings of 1st Workshop on Cognitive Models of Semantic Processing, University of Edinburgh, Edinburgh, Scotland.

Goro, T. and S. Akiba (2004a) The acquisition of disjunction and positive polarity in Japanese. In WCCFL 23: Proceedings of the 23rd West Coast Conference on Formal Linguistics. eds. V. Chand, A. Kelleher, A. J. Rodríguez, and B. Schmeiser, 251-264, Somerville, MA: Cascadilla Press.

Goro, Takuya (2007). Language-Specific Constraints on Scope Interpretation in First Language Acquisition. PhD dissertation, University of Maryland.

Grodzinsky, Y. and T. Reinhart (1993) "The Innateness of Binding and Coreference," Linguistic Inquiry 24, 69–102.

Gualmini, Andrea, Sarah Hulsey, Valentine Hacquard & Danny Fox. 2008. The Question–Answer Requirement for scope assignment. *Natural Language Semantics* 16(3). 205–237.

Gualmini, Andrea, Simona Maciukaite & Stephen Crain. 2003. Children's Insensitivity to Contrastive Stress in Sentences with ONLY. In Sudha Arunachalam, Elsie Kaiser & Alexander Williams (eds.), *Proceedings of the 25th Annual Penn Linguistics Colloquium*, 87-110. Philadelphia: University of Pennsylvania.

Halbert, A., S. Crain, D. Shankweiler, and E. Woodams (1995) "Children's Interpretive Use of Emphatic Stress," paper presented at the 8th Annual CUNY Conference on Human Sentence Processing, Tucson, Arizona.

Hirschbühler, P. (1982). VP deletion and across-the-board quantifier scope. In J. Pustejovsky & P. Sells (Eds.), Proceedings of NELS 12 (pp. 132-). Amherst: GLSA.

Huang, Yi Ting, Elizabeth Spelke & Jesse Snedeker. 2013. What Exactly do Numbers Mean? Language Learning and Development 9(2). 105–129.

Ioup, G. (1975). Some universals for quantifier scope. In J. P. Kimball (Ed.), Syntax and Semantics 4 (pp. 37-58). New York: Academic Press.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. Psychological Review, 87(4), 329-354.

Katsos, Napoleon & Dorothy V. M. Bishop. 2011. Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition* 120(1). 67–81.

Kurtzman, H. S., & MacDonald, M. C. (1993). Resolution of quantifier scope ambiguities. Cognition, 48, 243-279.

Leddon, E., J. Lidz & J. Pierrehumbert (2004). Suprasegmental cues to meaning in child-directed speech.17th CUNY Sentence Processing Conference, University of Maryland.

Lidz, J. and J. Musolino. 2002. Children's command of quantification. Cognition 84, 113-154.

MacDonald, M. C. (1994). Probabilistic Constraints and Syntactic Ambiguity Resolution. Language and Cognitive Processes, 9(2), 157-201.

Merchant, J. 2000. Antecedent-contained deletion in negative polarity items. Syntax 3, 144-150.

Micham, D. L., Catlin, J., VanDerveer, N. J., & Loveland, K. A. (1980). Lexical and Structural Cues to Quantifier Scope Relations. Journal of Psycholinguistic Research, 9(4), 367-377.

Musolino, J. (1998) Universal Grammar and the Acquisition of Semantic Knowledge: an Experimental Investigation into the Acquisition of Quantifier-Negation Interaction in English .Ph. D. Dissertation, University of Maryland at College Park.

Musolino, J., S. Crain and R. Thornton (2000) Navigating negative quantificational space. Linguistics 38-1, 1-32.

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

Musolino, J., and J. Lidz (2006) Why children aren't universaly successful with quantification. Linguistics 44: 817-852.

Neeleman, A. and T. Reinhart (1998) "Scrambling and the PF Interface," in W. Geuder and M. Butt, eds., Projecting From the Lexicon CSLI, Stanford University Press, Stanford, California.

Noveck, I. (2001) When children are more logical than adults: Experimental investigations on scalar implicatures. Cognition 78: 165-188.

Papafragou, A., and J. Musolino (2003) Scalar implicatures at the semantics/pragmatics interface. Cognition 80: 253-282.

Pinker, S. (1984) Language Learnability and Language Development. Cambridge, MA: MIT Press.

Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. Linguistics and Philosophy, 20, 335-397.

Reinhart, T. (1999) "The Processing Cost of Reference Set Computation: Guess Patterns in Acquisition," OTS Working Papers in Linguistics, 99–001–CL/TL, Utrecht University, Utrecht, The Netherlands.

Reinhart, Tanya. 2004. The Processing Cost of Reference Set Computation: Acquisition of Stress Shift and Focus. *Language Acquisition* 12(2). 109–155.

Reinhart, Tanya. 2006. *Interface Strategies: Optimal and Costly Computations*. Cambridge, MA: MIT Press.

Steedman, M. (2000). The syntactic process. Cambridge, Mass.: MIT Press.

Syrett, Kristen & Jeffrey Lidz. 2004. Children want to access every interpretation adults do. In Leah Bateman & Cherlon Ussery (eds.), Proceedings of the 35th Annual Meeting of the North East Linguistic Society [NELS], 591–602. Charleston, SC: Booksurge Publishing.

Syrett, Kristen & Simon, Georgia & Nisula, Kirsten. (2014). Prosodic disambiguation of scopally ambiguous quantificational sentences in a discourse context. Journal of Linguistics. 50. 453-493. 10.1017/S0022226714000012.

Szendrői, Kriszta. 2004. Acquisition evidence for an interface theory of focus. In Jacqueline van Kampen & Sergio Baauw (eds.), *Proceedings of Generative Approaches to Language Acquisition 2003*, 457–468. Utrecht: LOT.

Tunstall, S. 1998. The Interpretation of Quantifiers: Semantics and Processing. Doctoral dissertation, University of Massachusetts, Amherst.

Zhou, P., & Crain, S. (2009). Scope assignment in child language: Evidence from the acquisition of Chinese. Lingua, 119, 973–988.