ELSEVIER

Children's comprehension of sentences with focus particles

Kevin B. Paterson[a,*], Simon P. Liversedge[b], Caroline Rowland[c], Ruth Filik[d]

[a]*Department of Psychology, University of Leicester, Leicester, LE1 7RH, UK*
[b]*Department of Psychology, University of Durham, Durham, UK*
[c]*Department of Psychology, University of Liverpool, Liverpool, UK*
[d]*Applied Vision Research Unit, University of Derby, Derby, UK*

Abstract

We report three studies investigating children's and adults' comprehension of sentences containing the focus particle *only*. In Experiments 1 and 2, four groups of participants (6–7 years, 8–10 years, 11–12 years and adult) compared sentences with *only* in different syntactic positions against pictures that matched or mismatched events described by the sentence. Contrary to previous findings (Crain, S., Ni, W., & Conway, L. (1994). Learning, parsing and modularity. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum; Philip, W., & Lynch, E. (1999). Felicity, relevance, and acquisition of the grammar of *every* and *only*. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University conference on language development*. Somerville, MA: Cascadilla Press) we found that young children predominantly made errors by failing to process contrast information rather than errors in which they failed to use syntactic information to restrict the scope of the particle. Experiment 3 replicated these findings with pre-schoolers.
© 2003 Elsevier B.V. All rights reserved.

Keywords: Focus particles; Semantic interpretation; Language acquisition

1. Introduction

There exist in English a large number of expressions that are used to quantify over people, objects, events, locations, and time. These provide information about which

* Corresponding author.

entities and what number or proportion of them contribute to the meaning of a sentence. Within the psychological literature there has been considerable interest in how children reason with quantifiers as a means of assessing their logical competence (Braine & Rumain, 1983; Donaldson & Lloyd, 1974; Freeman & Stedmon, 1986; Hanlon, 1986; Inhelder & Piaget, 1964; Neimark & Chapman, 1975). However, the present experiments were concerned with fundamental issues about how children acquire syntactic cues to the meaning of quantifying expressions and learn to interpret sentences that contain them. Knowing how children learn to interpret sentences is an important precursor to knowing how they learn to reason with them.

Our studies were concerned with the focus particle *only*. Focus particles are a particular type of quantifying expression that are used to indicate contrastive focus (Rooth, 1992). English includes the focus particles *only*, *just*, *even*, *also* and *too* that when included in a sentence indicate that the extension of some linguistic element is to be contrasted with a set of alternatives. Some researchers have used the contrastive function of *only* to investigate whether non-syntactic information can guide the processing of syntactic ambiguities (Clifton, Bock, & Rado, 2000; Liversedge, Paterson, & Clayes, 2002; Ni, Crain, & Shankweiler, 1996; Paterson, Liversedge, & Underwood, 1999; Sedivy, 2002). Others have used it to investigate the acquisition and development of linguistic constraints on sentence interpretation (Crain, Ni, & Conway, 1994; Crain, Philip, Drozd, Roeper, & Matsuoka, 1992; Halbert, Crain, Shankweiler, & Woodams, 1995; Philip & Lynch, 1999). With the present studies we used the contrastive function of *only* to investigate children's use of syntactic cues to the meaning of quantifying expressions.

Part of the process of interpreting a sentence involves the reader or listener constructing a discourse model, which is a mental representation of persons, objects, relations and events described by the sentence, and other information that is inferred or specified by discourse context (e.g. Garrod & Sanford, 1994; Johnson-Laird, 1983). The discourse model for a sentence without a focus particle or other contrast-marking expression (e.g. Sanford, Moxey, & Paterson, 1996) must include representations of entities, relations and events that are made explicit in the text. Consider sentence (1):

1. Mary spoke with John.

The discourse model for this sentence will include representations of Mary and John, both participating in a speaking event at a previous point in time. Now consider the discourse model for a sentence that includes *only* but otherwise is identical to (1):

2. Only Mary spoke with John.

For this sentence *only* indicates contrastive focus and the resulting discourse model contrasts a set of entities that is made explicit and is the psychological subject of the sentence (i.e. the focus set) with some alternatives. As before, the discourse model will include representations of Mary and John, with Mary represented as the focus set and assigned the property of having spoken to John. In addition, it will include an alternative set (of unspecified persons) that did not speak with John. As this set is not made explicit readers may infer one using pragmatic knowledge or knowledge about the referential

context (e.g. Frazier, 1999; Sedivy, 2002). For instance, if (2) was preceded by the sentence *Mary and Louise went to the bar* then readers might infer that Louise did not speak to John.

Lexical, syntactic and prosodic cues determine the nature of the sets that are included in the discourse model. Lexical properties of the focus particle specify the contrast that is to be represented (e.g. Konig, 1991). Some particles, such as *only* and *just*, indicate an exclusive contrast, whereby a property that is specified as being true of the focus set is understood as being false for the alternative set. Others, such as *even* and *also*, indicate an additive contrast, whereby a property that is true of the alternative set focus set also is true for the focus set. Syntactic cues indicate which linguistic constituent denotes the focus set. Consider sentence (2). For this sentence the focus particle precedes the subject noun-phrase (i.e. it occurs in a pre-subject position) and indicates that a contrast is to be made between Mary and some other set (of persons). However, *only* can occur in other syntactic positions, with a change in position often indicating that a contrast is to be established for a different constituent (Konig, 1991). In sentence (3) it precedes the verb-phrase (a pre-verbal position) and indicates that a contrast should be established for one of its constituents.

3. Mary only spoke with John.

The focus particle is said to have different scope in sentences (2) and (3), where scope describes the linguistic domain from which function words, including focus particles, take their arguments. The scope of *only* is restricted to those linguistic constituents that it c-commands in the parse tree (Crain et al., 1994; Reinhart, 1983). When *only* occupies a pre-subject position it must take as its argument elements of the subject noun-phrase and readers contrast the focus set denoted by those elements with some alternatives. When it occupies a pre-verbal position it must take as its argument elements of the verb-phrase, and in this case readers contrast a focus set specified by these elements with a set of alternatives. Thus, scope restrictions require that when *only* occurs in pre-subject and pre-verbal positions the reader or listener constructs qualitatively different discourse representations for otherwise identical sentences.

Although syntactic cues restrict focus particles to taking scope over constituents within a certain linguistic domain, the reader or listener often must use knowledge about focus structure to determine which particular element denotes the focus set (Rooth, 1992). A standard assumption is that the potential scope of *only* is the linguistic domain specified by c-command rules, and the particular scope of the focus particle is the focussed linguistic constituent within this domain (Reinhart, 1999). Focus refers to the partitioning of a sentence into those constituents that are made prominent, or are the subject of psychological attention, and others that are in the background or subject to less attention (e.g. Chafe, 1976; Clark & Clark, 1973; Jackendoff, 1972). Focus can be conveyed by placing stress on an expression, e.g. *Only the old WOMAN laughed* (with capitals indicating the application of stress), or marked syntactically, such as by using cleft-constructions, e.g. *It was Mary who laughed* (e.g. Rooth, 1996). However, in written language focus often is ambiguous and readers must rely on processing preferences or contextual knowledge to determine the intended meaning of a sentence (Frazier, 1999).

Sentence (3) is an ambiguous sentence that can be interpreted with *only* taking restricted scope over constituents of the verb-phrase or taking unrestricted scope over the verb-phrase as a whole.[1] With scope restricted to the direct object the sentence means that John was the only person with whom Mary spoke, and with scope restricted to the verb it means that Mary spoke to John but did not do anything else with him. With unrestricted scope it means that the only thing Mary did was speak to John. Importantly, the possible interpretations of a sentence with restricted scope are a subset of the possible interpretations for the same sentence with unrestricted scope. Thus, if the only thing Mary did was speak to John then it also is true that John was the only person with whom she spoke. We return to this later.

At this point it should be clear that sentences with focus particles require the construction of more complex representations than do counterpart sentences without focus particles. The reader or listener must analyze the scope of the particle and the focus structure of the sentence to establish which linguistic constituent denotes the focus set. In addition, they must determine the nature of the contrast that is to be represented and use discourse context or pragmatic knowledge to infer a set of alternatives. With the present studies we examined two specific questions concerning the acquisition of these processing skills. First, we investigated whether there were developmental differences in the processing of contrastive focus for sentences with *only*. That is, do adults and children process sentences containing *only* as directing the construction of a discourse model that necessarily includes a focus and contrast set, for which a property that is true of the focus set is represented as false of the contrast set? Second, we examined whether children and adults made use of knowledge about the surface position of the particle to restrict its scope. That is, do children and adults process sentences with *only* in different syntactic positions as directing the construction of qualitatively different discourse models? Answers to these questions will make an important contribution to more general considerations of how children process scope ambiguities (e.g. Lidz & Musolino, 2002; Musolino, in press; Musolino, Crain, & Thornton, 2000).

Most psycholinguistic research into adult comprehension of quantification has been concerned with the focussing effects of quantifiers (e.g. Moxey & Sanford, 1987; Paterson, Sanford, Moxey, & Dawydiak, 1998; Sanford et al., 1996) and scope ambiguity resolution (e.g. Fodor, 1982; Frazier, Pacht, & Rayner, 1999; Ioup, 1975; Johnson-Laird, 1969; Kurtzman & MacDonald, 1993; Micham, Catlin, Van Derveer, & Loveland, 1980). Developmental research has included work on the acquisition of universal quantifiers and some limited work on the acquisition of *only* (e.g. Brooks & Braine, 1996; Bucci, 1978; Crain, 2000; Crain et al., 1994, 1996; Donaldson & Lloyd, 1974; Drozd, 2001; Drozd & van Loosbroek, 1998; Philip, 1995; Philip & Lynch, 1999; Philip & Takahashi, 1991). Research on universal quantifiers addresses similar problems to the ones considered here, and provides a framework for studying focus particles.

Studies have shown that young children do not interpret sentences with universal quantifiers in the same manner that adults do. When children evaluate a sentence like *Every man is carrying a sofa* they often judge it to be a false description of pictures in

---

[1] This is an example ambiguous sentence with pre-verbal *only*. Scope ambiguities may occur for other structures, including sentences with pre-subject *only*.

which all of the men are carrying a sofa but, in addition, there is an extra sofa. Adults are less likely to make such errors. This is a robust finding that can be traced back to Inhelder and Piaget (1964) and has been replicated in a variety of languages (Philip, 1995), although Crain et al. (1996) argued that the results are a consequence of the methods used to test children's comprehension.[2] The phenomenon often is referred to as "quantifier spreading", "symmetrical reading" or "conversion". Some accounts explain it in terms of children lacking syntactic knowledge about how to restrict the scope of quantifiers to the noun-phrase they modify (e.g. Bucci, 1978). Philip (1995) (also Philip & Takahashi, 1991) proposed that children instead process *every* as quantifying over the entire situation rather than specific noun-phrases. Thus, children treat *Every man is carrying a sofa* as specifying sets of events in which a man is carrying a sofa, and interpret the quantifier as requiring that for each occurrence of a man or sofa, the man is carrying the sofa. This implies that processing strategies undergo qualitative changes in the transition from child to adult performance.

Other accounts explain children's errors in terms of learning the mapping between linguistic form and the possible representations of a sentence (e.g. Brooks & Braine, 1996; Geurts, 2001). Brooks and Braine found that children below 10 years made errors for sentences with *each* but not for sentences with *all*. The results for *all* showed that children can use syntax to restrict the scope of universals. Brooks and Braine argued that children made errors for *each* as a consequence of learning how to map quantifiers onto semantic representations (i.e. discourse models).

They noted that universal quantifiers are associated with collective and distributive representations (e.g. Link, 1983). With a collective representation the modified phrase is assigned a group reading, thus *Every man carried a sofa* is represented as a group of men carrying a single sofa. A distributive representation requires a one-to-one correspondence, such that each man carried a different sofa. Brooks and Braine (1996) argued that from an early age children show the adult preference for associating *all* with collective or group representations, but take longer to acquire the adult preference for associating *each* with a non-exhaustive distributive representation. Non-exhaustive representations may include extra entities that are not part of the one-to-one correspondence. Exhaustive representations require that there are no extra entities. Brooks and Braine proposed that children make errors because they assign a distributive interpretation to sentences with *each* (and perhaps *every*) but adopt an exhaustive rather than non-exhaustive analysis.

Although the Geurts (2001) account differs in detail from that proposed by Brooks and Braine (1996), both require that children possess adult knowledge of syntactic structure, but lack adult knowledge about how to map sentences onto semantic representations.

---

[2] Crain et al. (1996) (see also Crain & Thornton, 1998) claimed that quantifier spreading effects occur as a result of the methods used in these studies. In the standard task the experimenter considers cases in which the sentence should be judged to be true. Participants view pictures in which there is a one-to-one correspondence between entities (e.g. equal numbers of men and sofas) or a one-to-one correspondence plus extra instances of the unmodified term (e.g. extra sofas). Crain et al. argued that not including cases in which the sentence is demonstrably false (e.g. by including in the picture extra men who are not carrying a sofa) might bias children towards making a false response to the picture with extra sofas. They claimed that including false cases eliminated errors, however Philip and Lynch (1999) obtained the standard pattern of results despite modifying the task to eliminate any response bias.

Geurts' account is based on formal linguistic approaches to natural language quantification (e.g. Barwise & Cooper, 1981), and distinguishes between "strong" and "weak" quantifiers. Strong quantifiers such as *each*, *every*, *all*, and *most* presuppose a set in the discourse context and indicate what proportion of that set has a particular property. For instance, *Most men carried a sofa* presupposes a set of men, the majority of whom carried a sofa. Weak quantifiers (e.g. *some*) do not carry proportional presuppositions. Geurts argued that children at first assign a weak reading to all quantifiers, and because they are missing presuppositional information children may misinterpret universals as ranging over the wrong domain. Importantly, this account, and the one proposed by Brooks and Braine, do not require that children make errors because they lack syntactic knowledge or employ non-adult strategies, but because they need to fine-tune the mapping between quantifiers and discourse models. With the current studies we draw a similar contrast between accounts explaining children's errors with focus particles as due to them lacking syntactic restrictions on scope and an alternative account that explains those errors in terms of the mapping between linguistic form and the discourse model.

While there has been little previous research into the acquisition of focus particles, Crain et al. (1994) did propose a detailed account of their acquisition that is integrated with an account of adult sentence processing. They argued that whereas adults process sentences using a combination of syntactic rules and principles originally outlined within the Referential theory of sentence processing (e.g. Crain & Steedman, 1985), young children employ qualitatively different strategies that enable them to acquire the possible analyses of a sentence. Critically, Crain et al. proposed that young children do not use syntax to restrict the scope of focus particles. Whereas adults process (2) with scope restricted to the subject noun-phrase, and process (3) with *only* taking scope over constituents of the verb-phrase, young children process both sentences in terms of a single representation, regardless of where the particle occurs. As children encounter situations in which the analysis they initially apply proves incorrect, they accumulate evidence for alternative analyses, and once they have accumulated sufficient evidence they will assign distinct discourse representations to sentences with *only* occurring in different surface positions.

Adults are thought to employ two general principles when resolving scope ambiguities. The Principle of Referential Success stipulates that when an ambiguous sentence occurs within a disambiguating referential context, adults favour an interpretation that refers to entities from that context. Thus, adult scope decisions are guided by the availability of plausible focus and contrast sets in the discourse context. The Principle of Parsimony stipulates that an ambiguous sentence is assigned the analysis that requires fewest inferences (presuppositions) about information that is not made explicit. The focus set is always explicit because the focus particle necessarily takes scope over the linguistic constituent that denotes it. Therefore, decisions about which analysis to adopt depend on the relative number of inferences needed to establish a contrast set. Crain et al. claimed that it is most parsimonious to instantiate the contrast set with fewest members, and that adults preferentially adopt a restricted analysis for this reason. Consider sentence (3). With *only* taking scope over the direct object the contrast set includes events in which Mary spoke with someone other than John. With *only* taking unrestricted scope it includes events in which Mary did something other than speak with John. Because speaking with

someone else is a subset of doing something else, the restricted scope analysis presupposes a smaller contrast set than does the corresponding unrestricted scope analysis.

Children do not receive explicit instruction or feedback about how to interpret ambiguous sentences and must employ strategies that enable them to acquire semantic interpretations without instruction. Therefore, Crain et al. proposed that children acquire alternative scope analyses by employing the Semantic Subset Principle. It stipulates that children preferentially adopt the analysis that is most specific, and as they encounter situations in which it proves to be incorrect, they accumulate evidence for the use of alternative analyses. Following this principle, children should preferentially assign an unrestricted analysis to an ambiguous sentence. Consider sentence (3). With scope restricted to the direct object the sentence is true if Mary speaks to no one other than John; its unrestricted analysis is true if she does *nothing* other than speak with John. Not speaking with anyone else is a subset of not doing anything else therefore the unrestricted analysis is favoured by the Semantic Subset Principle because it is true in the most specific, or narrowest, set of circumstances.

Several studies appear to demonstrate that children assign a single representation to sentences with *only* regardless of its surface position (Crain et al., 1992; Drozd & van Loosbroek, 1998; Philip & Lynch, 1999). First, Crain et al. examined how 3- to 6-year-old children compared sentences like (4) and (5) with a picture of a cat holding a flag, a duck holding a flag and a balloon, and a frog holding a balloon.

4. Only the cat is holding a flag.
5. The cat is only holding a flag.

If participants used syntax to restrict scope then they should have judged (5) to be a true description of depicted events and judged (4) to be false. However, most participants judged both sentences to be true. Crain et al. took this to indicate that they adopted an analysis with *only* taking scope over the verb-phrase, regardless of its surface position. In a subsequent study, Philip and Lynch (1999) (also Drozd & van Loosbroek, 1998) examined how adults and pre-schoolers compared a sentence with pre-subject *only*, e.g. *Only the dog is holding an octopus*, against a picture of a dog holding an octopus and a starfish, and two cats holding nothing. Adults judged the sentence to be a true description of events, whereas over a third of pre-schoolers judged it to be false. Thus, it appeared that some children misanalyzed the sentence with *only* taking scope over the direct object rather than the subject noun-phrase.

These studies produced striking results but we have concerns about the methodologies that lead us to question whether the results necessarily support the experimenters' conclusions. In the Crain et al. study, children compared sentences with pre-subject and pre-verbal *only* against a picture depicting events that were consistent with verb-phrase scope and inconsistent with subject noun-phrase scope. However, the events also were consistent with the analysis that would be assigned to counterpart sentences without *only*, e.g. *The cat is holding a flag*. Thus, although there was evidence that children did not process target sentences with *only* taking scope over the subject noun-phrase, the study did not unequivocally show that children instead processed them with *only* taking scope over the verb-phrase.

There is at least one alternative explanation of the results obtained by Crain et al. Perhaps children did not always interpret sentences by constructing a discourse model in which a focus set was contrasted with an explicit set of alternatives. Instead, they may have constructed a discourse model containing only the information given in the sentence. That is to say, they might have instantiated a focus set comprising a cat holding a flag without computing a contrast. If this were the case, then they may have interpreted sentences with pre-verbal and pre-subject *only* as having the same meaning as counterparts without *only*. Importantly, if children do not always include explicit sets of alternatives in their discourse model for sentences with *only*, then they may judge sentences to be true of any state of affairs that are consistent with events represented by the focus set, regardless of where *only* occurs in the sentence. A thorough test of children's use of syntactic information when processing sentences with *only* must include conditions that enable the experimenter to determine if participants use information about both focus and alternative sets. We also were concerned that Crain et al. did not examine adult performance in their experimental task. Although other studies (e.g. Philip & Lynch, 1999) have shown that adults use syntax to restrict the scope of focus particles, Crain et al. did not do so in their experiment. In order to be certain that results obtained by Crain et al. demonstrate that children differ from adults by not attending to the surface position of *only*, it is necessary to compare child and adult performance on the same task.

Flaws in the Philip and Lynch (1999) study also undermined our confidence in their conclusions. They used a complex task in which participants first compared the target sentence with two other pictures before making the critical comparison. Sentence–picture comparisons were made in a fixed order as part of a story. First the sentence was compared with a picture of the dog holding an octopus and a cat holding nothing, then with one of the dog holding an octopus and a starfish, and a cat holding a octopus. Philip and Lynch only reported data for participants who responded correctly to the first two pictures (but without reporting error rates). It would be unsurprising if pre-schoolers found this task confusing, and its complexity may well have impaired their performance. However, the study lacked control conditions that were needed to show that pre-schoolers performed the task satisfactorily. Had Philip and Lynch included sentences without *only* and ones with pre-verbal *only* then it might have been possible to demonstrate that pre-schoolers made errors exclusively for sentences with pre-subject *only*. This would have given greater credence to the claim that the errors were due to pre-schoolers misanalyzing scope for these sentences.

To address our concerns we conducted three experiments in which we tested whether adults and children interpreted sentences with pre-subject and pre-verbal *only* by constructing discourse models that include focus and alternatives sets, and whether they used syntax to restrict scope. We compared the processing of simple sentences with pre-verbal *only* (e.g. 6) and pre-subject *only* (e.g. 7), against counterparts without *only* (e.g. 8).

6. The fireman is only holding a hose.
7. Only the fireman is holding a hose.
8. The fireman is holding a hose.

2. Experiment 1

In Experiment 1 we examined performance of three groups of children (6- to 7-year-olds, 8- to 10-year-olds and 11- to 12-year-olds) and adults on a forced-choice picture-selection task. Participants indicated if sentences were true or false descriptions of events depicted in each of six pictures. If participants interpret sentences with *only* by instantiating focus and alternative sets in their discourse model then they should select a different combination of pictures for sentences with *only* than without *only*. However, if participants fail to instantiate appropriate alternative sets in their discourse model they should select the same combination of pictures for sentences with and without *only*. Second, if participants use c-command rules to restrict the scope of the focus particle, then they should judge a different combination of pictures to be consistent with events described by sentences with pre-subject and pre-verbal *only*. Conversely, if participants do not use information about the surface position of the focus particle to restrict its scope then they should select the same combination of pictures for sentences with pre-subject and pre-verbal *only*.

From the Crain et al. study we cannot determine the transition age at which children should begin to process sentences in the same manner as adults. However, following their study we would expect a different pattern of results for adults than we would expect for the youngest children. According to Crain et al.'s account, young children should assign the same scope analysis to sentences with pre-subject and pre-verbal *only*, regardless of the surface position of the focus particle. In contrast, we predict that young children often might fail to instantiate an explicit alternative set as part of their discourse model for sentences with *only*. This may occur because children fail to attend to the focus particle, or fail to construct appropriate discourse representations. Consequently, they predominantly should make errors in which they judge sentences with *only* as having the same meaning as counterparts without *only*.

Although Crain et al. are not clear about precisely when in the course of language development children will begin to process sentences with *only* in the same manner as adults, there are two possibilities. First, children may begin to process sentences in the same way as adults during a critical period of language development, in which case we would expect their error rate to decrease markedly at a particular age to a level that is comparable with adults. Alternatively, children may gradually approximate adult performance in the course of development in which case we would expect the error rate gradually to reduce across age groups. Our studies will enable us to determine which is the case.

In our experiment, participants viewed and were read the target sentence and compared it with each of six pictures depicting events that matched or mismatched the sentence. Example pictures for sentences (6)–(8) are shown in Fig. 1.

Each picture depicts two characters: the character from the sentence (e.g. a fireman) and a secondary character that is not explicitly mentioned (e.g. a policeman). Fig. 1A always matches with the sentence as the main character performs the action from the sentence (e.g. he is holding a hose) and the secondary character does not, whereas Fig. 1B always mismatches with the sentence, as the secondary character performs described action, and the main character does not. Fig. 1C–E discriminate between alternative scope analyses. In Fig. 1C, both characters perform the action from the sentence. In Fig. 1D, the main character performs both the action from the sentence and a contrast action (e.g. holding
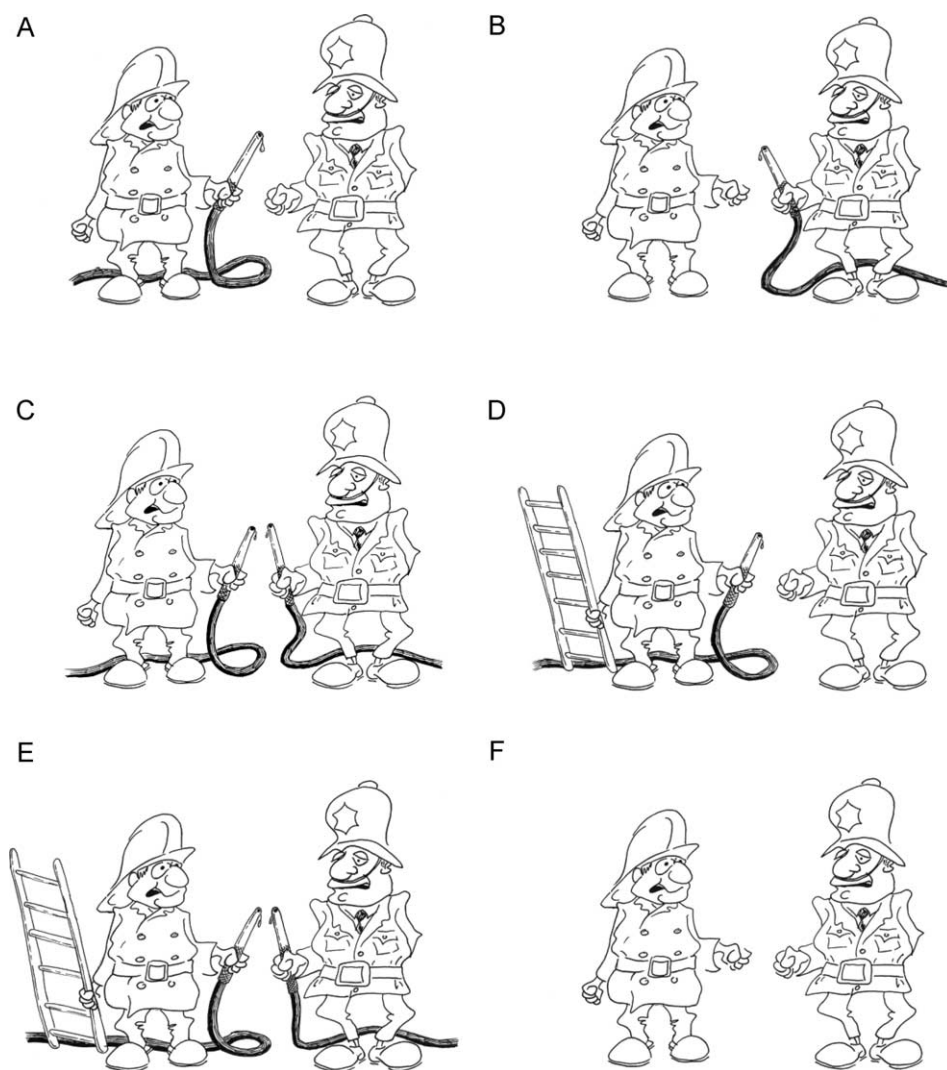
Fig. 1. Examples of cartoon drawings used in Experiments 1, 2 and 3.

a ladder), while the secondary character does nothing. In Fig. 1E, the main character performs the action from the sentence and the contrast action, while the secondary character performs the action from the sentence. Fig. 1C–E match sentences that do not include a focus particle, as the main character always performs the action described in the sentence. Fig. 1C matches with an interpretation in which the focus particle takes scope over the direct object, and mismatches with one where it takes scope over the subject noun-phrase. Conversely, Fig. 1D matches with an analysis where the focus particle takes scope over the subject noun-phrase, and mismatches with one where the particle takes scope over the direct object. Fig. 1E mismatches with both scope interpretations. Finally,

Fig. 1F always mismatches with the sentence, as neither character performs the described action.

From these materials we generated the following predictions. For sentences without *only*, participants should judge the sentence to be a true description of Fig. 1A,C–E, and a false description of Fig. 1B,F. If sentences with *only* are processed with the focus particle taking scope over the subject noun-phrase, then they should be judged to be a true description of Fig. 1A,D, and a false description of Fig. 1B,C,E,F. If sentences with *only* are processed with the focus particle taking scope over the direct object, then they should be judged to be a true description of Fig. 1A,C, and false of Fig. 1B,D–F. Finally, if sentences with *only* are processed without including explicit sets of alternatives as part of the discourse model, then they will generate the same pattern of responses as obtained for counterpart sentences without *only*.

2.1. Method

2.1.1. Participants

Ninety-seven English speaking participants were divided into four age-groups: 25 6- to 7-year-olds and 22 8- to 10-year-olds from Long Clawson Church of England Primary School in Melton Mowbray, 25 11- to 12-year-olds from Long Field High School in Melton Mowbray, and 25 undergraduates at the University of Derby.

2.1.2. Materials and design

We constructed 12 sets of present tense simple sentences that are listed in Appendix A. One version of each sentence did not contain *only* (e.g. 6). In the second version, *only* occurred in a pre-subject position (e.g. 7), and in the third version, *only* appeared in the pre-verbal position (e.g. 8). For each set of sentences, we prepared a set of six cartoon pictures, with example pictures in Fig. 1.

There were two independent variables: age group and sentence type. The dependent variable was the frequency of different types of response: responses that matched the analysis of sentences without *only*, responses with scope restricted to the subject noun-phrase, responses with scope restricted to the direct object, and other responses. Sentences were divided into three lists, with one version of each sentence in each list. Each participant viewed one list. Participants viewed four sentences of each type and no participant viewed more than one version of each sentence.[3]

2.1.3. Procedure

Participants were tested individually, and each test session lasted approximately 20 minutes. Each participant viewed sentences from one list, and compared each sentence against the corresponding set of six pictures. Each sentence was written on a card and read

---

[3] An error was made in data collection for 10- to 12-year-old participants in Experiment 1. In this condition participants viewed all versions of sentence items and compared these against the corresponding set of pictures. Consequently, participants in this condition viewed 12 sentences without *only*, 12 with pre-verbal *only* and 12 with pre-subject *only*. Although it is likely that this error in data collection will have cued participants to the nature of the study (as they viewed the same sentence in all experimental conditions) it did not unduly influence results.

aloud to the subject. Participants were shown pictures one at a time and asked to indicate if the sentence was a true or false description of the depicted events. Participants responded 'yes' if they thought the sentence described events in the picture, and responded 'no' if they thought otherwise. Sentences were presented in a fixed order. For half the participants, pictures were presented in order A to F, and for the other half they were presented in order F to A.

2.2. Results and discussion

We first categorized participants' responses as belonging to one of four categories. A response was coded as *no scope* analysis if participants judged a sentence to be true when the main character performed the action from the sentence (i.e. Fig. 1A,C–E) and false otherwise (i.e. Fig. 1B,F). This corresponded to the correct interpretation of sentences without *only*. Alternatively, responses were coded as having scope restricted to the subject noun-phrase (*NP scope*) if sentences were judged to be true when the main character but not the secondary character performed the described action (i.e. Fig. 1A,D) and false otherwise (i.e. drawings Fig. 1B,C,E,F). Responses were coded as having scope restricted to the direct object (*VP scope*) if sentences were judged to be true when the main character performed the described action but not the contrast action (i.e. Fig. 1A,C), and false otherwise (i.e. Fig. 1B,D–F). A fourth category was included for other types of response. Table 1 shows the mean percentage frequency (with standard deviations) of responses across age groups and sentence type.

Table 1
Mean percentage responses (with standard deviations) for participants in each age group when evaluating sentences without *only*, with pre-subject *only* and with pre-verbal *only* in Experiment 1

Sentence	Age	No scope	NP scope	VP scope	Other
Without *only* (no scope = correct response)	Adult	88.0 (23.0)	1.0 (5.0)	4.0 (11.8)	7.0 (13.5)
	11–12 years	81.0 (25.3)	1.0 (5.0)	7.0 (11.5)	11.0 (21.7)
	8–10 years	67.0 (28.3)	21.0 (18.8)	1.0 (1.8)	11.0 (21.3)
	6–7 years	92.0 (15.7)	0.0 (0.0)	0.0 (0.0)	8.0 (15.7)
	All ages	82.0 (23.1)	5.8 (7.2)	3.0 (6.3)	9.3 (18.0)
Pre-subject *only* (NP scope = correct response)	Adult	1.0 (5.0)	81.0 (22.0)	0.0 (0.0)	18.0 (21.1)
	11–12 years	9.0 (17.5)	59.0 (37.4)	7.0 (10.2)	25.0 (22.7)
	8–10 years	11.0 (26.7)	55.0 (29.4)	7.0 (10.2)	27.0 (22.7)
	6–7 years	17.0 (23.6)	56.0 (38.6)	4.0 (11.8)	23.0 (22.7)
	All ages	9.5 (18.2)	62.8 (31.9)	4.5 (8.0)	23.3 (22.3)
Pre-verbal *only* (VP scope = correct response)	Adult	4.0 (11.8)	9.0 (18.9)	60.0 (33.9)	27.0 (24.9)
	11–12 years	11.0 (22.9)	10.0 (17.7)	60.0 (33.1)	19.0 (22.0)
	8–10 years	5.0 (21.3)	23.0 (28.6)	47.0 (34.6)	25.0 (22.3)
	6–7 years	21.0 (30.3)	9.0 (21.5)	44.0 (36.3)	26.0 (28.4)
	All ages	10.3 (21.6)	12.8 (21.7)	52.8 (34.5)	24.3 (24.4)

Most responses were coded as belonging to *no scope*, *NP scope* or *VP scope* categories. For the other types of response 1.2% were errors in which participants responded incorrectly to pictures that were always true of the sentence (Fig. 1A) or always false of the sentence (Fig. 1B,F). In the other cases participants made contradictory responses to pictures that represented contrast information (e.g. they judged the sentence to be true of Fig. 1A,C,D or Fig. 1A,E and false of the others).

We examined the frequency of correct responses for each sentence type (i.e. *no scope* analyses for sentences without *only*, *NP scope* analyses for sentences with pre-subject *only*, and *VP scope* analyses for sentences with pre-verbal *only*). Data were analyzed using two 4 (age group) × 3 (sentence type) mixed design ANOVAs, one treating subjects as a random variable (F_1) and one treating sentences as a random variable (F_2) (Clark, 1973). There was a significant main effect of sentence type ($F_1(2, 186) = 41.92$, $P < 0.001$; $F_2(2, 88) = 97.00$, $P < 0.001$). Participants produced more correct responses for sentences without *only* than ones with pre-subject *only* ($F_1(1, 93) = 56.59$, $P < 0.001$; $F_2(1, 44) = 113.65$, $P < 0.001$). Furthermore, there were more correct responses for sentences with pre-subject than pre-verbal *only* ($F_1(1, 93) = 6.64$, $P < 0.05$; $F_2(1, 44) = 8.64$, $P < 0.001$). The main effect of age group was not reliable ($F_1(3, 93) = 2.19$, $P > 0.05$; $F_2(3, 44) = 8.64$, $P < 0.001$).

There was a significant interaction of age group and sentence type ($F_1(6, 186) = 2.26$, $P < 0.05$; $F_2(6, 88) = 5.18$, $P < 0.01$). To examine this interaction further, we conducted a series of one-way between-subjects ANOVAs on the frequency of correct responses for each sentence type. There was no reliable difference in the frequency of correct responses across age groups for sentences without *only* ($F_1(3, 97) = 1.15$, $P > 0.05$; $F_2(3, 44) = 2.71$, $P < 0.06$), or with pre-verbal *only* ($F_1(3, 97) = 1.51$, $P > 0.05$; $F_2(2, 44) = 6.46$, $P < 0.01$). However, for pre-subject *only*, there was a significant effect of age group ($F_1(3, 97) = 3.65$, $P < 0.01$; $F_2(3, 44) = 9.04$, $P < 0.001$). Post-hoc tests indicated that adults produced more correct responses than either 6- to 7-year olds ($P < 0.05$) or 8- to 10-year-olds ($P < 0.05$).[4]

The high rate of correct responses across age groups for sentences without *only* indicated that participants could understand and perform the experimental task. Furthermore, the results indicated that 6- to 7-year-olds experienced more difficulty than did adults when evaluating sentences with pre-subject *only*. The error rate for these sentences decreased monotonically across age groups and there was no evidence for the sudden shift in error rate that would be expected if there was a critical period for the acquisition of appropriate discourse formations for sentences containing focus particles. Unexpectedly, all age groups of participants found it difficult to evaluate sentences with pre-verbal *only*. The error rate for these sentences is approximately equal across age groups.

We next examined responses produced by 6- to 7-year-olds and adults in more detail. This enabled us to examine Crain et al.'s claim that young children make errors because they fail to use syntactic rules to restrict the scope of focus particles,

---

[4] Post-hoc Games–Howell tests were used to adjust for unequal subject numbers across age groups (Howell, 1997). All significant differences obtained using these tests also were obtained using post-hoc Tukey tests.

Table 2
Mean percentage selection (with standard deviations) of each picture by adults and 6- to 7-year-olds when evaluating sentences without *only*, with pre-subject *only* and with pre-verbal *only* in Experiment 1

Age group	Sentence	Picture					
		A	B	C	D	E	F
Adults	Without *only*	100.0* (0.0)	1.0 (5.0)	96.0* (13.8)	91.0* (21.5)	92.0* (22.5)	0.0 (0.0)
	Pre-subject *only*	99.0* (5.0)	0.0 (0.0)	2.0 (6.9)	85.0* (21.7)	5.0 (10.2)	0.0 (0.0)
	Pre-verbal *only*	98.0* (6.9)	2.0 (6.9)	74.0* (26.5)	15.0 (23.9)	4.0 (9.4)	1.0 (5.0)
6–7 years	Without *only*	99.0* (5.0)	1.0 (5.0)	98.0* (6.9)	96.0* (11.8)	97.0* (8.3)	0.0 (0.0)
	Pre-subject *only*	100.0* (0.0)	1.0 (5.0)	41.0 (36.0)	88.0* (19.3)	37.0 (40.3)	0.0 (0.0)
	Pre-verbal *only*	100.0* (0.0)	0.0 (0.0)	91.0* (21.5)	37.0 (35.4)	36.0 (33.1)	0.0 (0.0)

Asterisks indicate pictures that match with sentence meaning.

and to compare it with our claim that errors predominately are due to children failing to instantiate explicit alternative sets. Table 2 shows the frequency with which adults and 6- to 7-year-olds judged each picture to be a true depiction of events for sentences without *only* and for sentences with and without *only*.

The frequency data further demonstrate that young children and adults produced a high rate of correct responses for sentences without *only*. Both judged sentences without *only* to be true descriptions of pictures depicting the main character performing the described action (i.e. Fig. 1A,C–E) with few responses in which they judged sentences to be true of pictures for which this was not the case.

Now consider sentences with *only*. Participants correctly evaluated ones with pre-subject *only* if they judged Fig. 1A,D to be true and the others to be false, and correctly evaluated sentences with pre-verbal *only* if they judged Fig. 1A,C to be true and the others to be false. If participants misanalyzed scope then we would expect them to make errors by judging Fig. 1C to be true for pre-subject *only* and judging Fig. 1D to be true for pre-verbal *only*. However, if participants failed to attend to contrast information then we would expect them erroneously to judge Fig. 1C–E to be true for sentences with *only*. Thus, the critical comparisons involved inspecting error rates for Fig. 1C–E. Note that responses to Fig. 1C,D alone do not discriminate between the alternative accounts. However, the critical findings are whether we obtained any errors to Fig. 1E. To be clear, errors made by judging Fig. 1E to be true for sentences with *only* are wholly consistent with our account, but are not predicted by the alternative Crain et al. account. Errors made by erroneously judging Fig. 1C or Fig. 1D to be true for sentences with *only* are predicted by both accounts.

An inspection of Table 2 indicates that 6- to 7-year-olds produced the pattern of responses predicted by our account, with participants in this age group judging Fig. 1A, C–E to be true for sentences with *only* on a large number of trials. Critically, young children judged Fig. 1E to be true for sentences with *only* on over one-third of trials. Such responses were consistent with our account but offered no support to Crain et al.'s account. There was comparatively little difference in the frequency of judging Fig. 1C or Fig. 1E to

be true for pre-subject *only*, or in judging Fig. 1D or Fig. 1E to be true for pre-verbal *only*. Adults produced a qualitatively different pattern of results for sentences with *only*, with fewer errors in which they erroneously judged Fig. 1C, Fig. 1D or Fig. 1E to be true for sentences with *only*. Thus, a visual inspection of the raw frequency data suggests that results are consistent with the claim that children often make errors by neglecting contrast information and that they differ from adults in this respect.

A more rigorous test of the competing hypotheses involves returning to the categorized data in Table 1, and comparing the frequency of responses in which adults and 6- to 7-year-olds unambiguously failed to attend to contrast information (i.e. *no scope* responses) with responses in which they unambiguously misanalyzed scope. Misanalyzing scope involved producing *VP scope* responses for pre-subject *only* and *NP scope* responses for pre-verbal *only*. Data were analyzed using two 2 (age group) × 2 (sentence type) × 2 (response type) ANOVAs.

We found that adults made fewer errors than 6- to 7-year-olds ($F_1(1, 48) = 11.03$, $P < 0.01$; $F_2(1, 22) = 36.60$, $P < 0.001$). The main effect of sentence type was significant by subjects and marginal by items ($F_1(1, 48) = 5.16$, $P < 0.05$; $F_2(1, 22) = 3.90$, $P < 0.07$), with more errors for pre-verbal than pre-subject *only*. The main effect of response type was not reliable ($F_1(1, 48) = 2.62$, $P > 0.05$; $F_2(1, 22) = 9.55$, $P < 0.01$), and there was no interaction of sentence and response type or sentence type and age group (Fs < 1.7). However, the interaction of age-group and response type was significant ($F_1(1, 48) = 5.00$, $P < 0.05$; $F_2(1, 22) = 17.00$, $P < 0.001$). Post-hoc tests indicated that 6- to 7-year-olds produced more *no scope* than scope analysis errors ($P < 0.01$) but there was no difference in the frequency of no scope and scope analysis errors for adults ($P > 0.05$). An inspection of Table 1 indicates that 6- to 7-year-olds made almost three times as many errors in which they neglected contrast information than errors in which they misanalyzed scope. There was no three-way interaction of age group, sentence and response type (Fs < 1).

In summary, children and adults correctly evaluated sentences without *only*, but all ages made a large proportion of errors for sentences with pre-verbal *only*. For sentences with pre-subject *only* young children made most errors and the error rate decreased across age groups, to the extent that an inspection of Table 1 suggests that adults found it as easy to correctly evaluate sentences with pre-subject *only* as sentences without *only*. Two sources of evidence suggested that young children often made errors by failing to attend to contrast information. First, young children frequently judged Fig. 1E to be true for sentences with *only*. Such responses were consistent with participants making errors by neglecting contrast information, but inconsistent with them making errors by misanalyzing scope. Furthermore, a formal analysis of categorized responses demonstrated that young children made significantly more errors by neglecting contrast information than errors of scope misanalysis, with no difference in the frequency of these types of error for adults.

Thus, our results indicated that adults used knowledge about focus and alternative sets when evaluating sentences with *only*, but that young children often made errors by failing to take account of information about the contrast set. The results are problematic for Crain et al.'s account of children's sentence interpretation, which requires that children make errors by failing to restrict the scope of the focus particle.

3. Experiment 2

The results we obtained in Experiment 1 lead us to conclusions that are quite different from those reported by Crain et al. (1992). Whereas they found that the majority of participants misanalyzed the scope of the focus particle in sentences containing *only*, we found that young children made more errors in which they misanalyzed sentences by neglecting contrast information. One reason why these differences in performance might have occurred is that the specific task employed in our study differed from that employed by Crain et al. Whereas participants in our study made a forced-choice decision for each picture in turn, participants in Crain et al.'s study simultaneously viewed all of the alternative pictures before selecting the one that was consistent with the test sentence. Perhaps when all the alternative pictures are viewed in parallel participants are provided with more of a cue to using contrast information. In order to rule out this possibility we conducted our study again, but this time participants were shown all six pictures simultaneously and asked to select those pictures that were consistent with the sentence.

3.1. Method

3.1.1. Participants
One hundred participants were divided into four age groups: 26 6- to 7-year-olds from Richard Alderman Primary School in Norfolk, 22 8- to 10-year-olds and 26 11- to 12-year-olds from Drayton Middle School in Norfolk, and 26 undergraduates from the University of Derby. No adult participant had also participated in Experiment 1.

3.1.2. Materials and design
We used the same sentences and pictures as used in Experiment 1. Sentences were divided into three lists, each list containing one version of each sentence and four of each type. There were two independent variables: age group of participants and sentence type. The dependent variable was the frequency of *no scope*, *NP scope*, *VP scope* and *other* types of response.

3.1.3. Procedure
Participants were tested individually, and each test session lasted about 20 minutes. Participants were shown and read aloud sentences from one of the three lists one at a time. Participants were then shown all six corresponding drawings at once and asked to indicate which of the drawings matched with the sentence.

3.2. Results and discussion

As with Experiment 1, we first coded participants' responses into *no scope*, *NP scope*, *VP scope* or *other* categories. Table 3 shows the mean percentage scores (with standard deviations) for each response category across age groups and sentence type.

As with Experiment 1 most responses belonged to *no scope*, *NP scope* or *VP scope* categories. For 0.3% of the other responses, participants responded incorrectly to Fig. 1A,

Table 3
Mean percentage responses (with standard deviations) for participants in each age group when evaluating sentences without *only*, with pre-subject *only* and with pre-verbal *only* in Experiment 2

Sentence	Age	No scope	NP scope	VP scope	Other
Without *only* (no scope = correct response)	Adult	93.3 (11.3)	0.0 (0.0)	1.0 (4.9)	5.8 (10.7)
	11–12 years	89.4 (30.9)	7.7 (17.0)	1.9 (6.8)	1.0 (4.9)
	8–10 years	73.9 (35.8)	3.4 (8.8)	5.8 (16.3)	12.5 (27.5)
	6–7 years	84.6 (30.9)	3.8 (15.3)	5.8 (16.3)	5.8 (20.4)
	All ages	85.3 (27.2)	3.7 (10.3)	3.6 (11.1)	6.3 (15.9)
Pre-subject *only* (NP scope = correct response)	Adult	1.0 (4.9)	89.4 (14.4)	1.9 (6.8)	7.7 (11.8)
	11–12 years	4.8 (12.3)	71.2 (33.7)	9.6 (24.6)	14.4 (18.9)
	8–10 years	55.7 (40.0)	19.3 (28.8)	17.0 (31.2)	10.2 (22.7)
	6–7 years	45.2 (41.8)	26.9 (33.1)	16.3 (28.2)	11.5 (20.3)
	All ages	26.7 (24.8)	51.7 (27.5)	11.2 (22.7)	11.0 (18.4)
Pre-verbal *only* (VP scope = correct response)	Adult	0.0 (0.0)	3.8 (9.2)	82.7 (27.2)	13.4 (21.5)
	11–12 years	6.7 (16.7)	18.3 (27.0)	65.4 (34.0)	9.6 (14.3)
	8–10 years	10.2 (22.7)	15.9 (30.4)	56.8 (41.7)	17.0 (27.1)
	6–7 years	49.0 (45.0)	4.8 (15.8)	34.6 (39.4)	11.5 (20.3)
	All ages	16.5 (21.1)	10.7 (20.6)	59.9 (35.6)	12.9 (20.8)

Fig. 1B or Fig. 1F, and for the remainder participants made a contradictory set of responses to pictures that represented contrast information.

We examined the frequency of correct responses for each sentence type using two 4 (age group) × 3 (sentence type) mixed design ANOVAs, one treating subjects as the random variable (F_1), and the other treating stimuli items as the random variable (F_2). There was a significant main effect of age group ($F_1(3, 96) = 20.36$, $P < 0.001$; $F_2(2, 88) = 23.49$, $P < 0.001$). Adults produced more correct responses than the other age groups ($P < 0.05$) and 11- to 12-year-olds produced more correct responses than 6- to 7-year-olds ($P < 0.05$). There was no difference in the frequency of correct responses for 6- to 7-year-olds and 8- to 10-year-olds ($P > 0.05$).

In addition, there was a significant main effect of sentence type ($F_1(2, 192) = 22.56$, $P < 0.001$; $F_2(2, 44) = 23.49$, $P < 0.001$), with more correct responses for sentences without *only* than for sentences with pre-subject *only* ($F_1(1, 96) = 41.23$, $P < 0.001$; $F_2(1, 44) = 31.82$, $P < 0.001$). Unlike Experiment 1 there was no difference in the frequency of correct responses for pre-subject and pre-verbal *only* (Fs < 1).

Finally, there was a significant interaction of age group and sentence type ($F_1(6, 192) = 4.13$, $P < 0.01$; $F_2(6, 88) = 4.82$, $P < 0.001$). In order to examine this interaction further, we conducted one-way between subjects ANOVAs, comparing the frequency of correct responses across age groups of participants for each sentence type. For sentences without *only* the effect of age group was not reliable ($F_1(3, 96) = 2.43$, $P > 0.05$; $F_2(3, 47) = 3.24$, $P < 0.05$). For sentences with pre-subject *only*, there was

a significant effect of age group ($F_1(3, 96) = 18.43$, $P < 0.001$; $F_2(3, 47) = 28.38$, $P < 0.001$). Post-hoc tests indicated that adults produced more correct responses than 8- to 10-year-olds and 6- to 7-year-olds ($P < 0.05$). The 11- to 12-year-olds and 8- to 10-year-olds also produced more correct responses than 6- to 7-year-olds ($P < 0.05$). For sentences with pre-verbal *only* there was a significant effect of age group ($F_1(3, 96) = 8.12$, $P < 0.001$; $F_2(3, 47) = 8.96$, $P < 0.001$). Adults and 11- to 12-year-olds produced more correct responses than 6- to 7-year-olds ($P < 0.05$).

As with Experiment 1, all four age groups found it easy to correctly evaluate sentences without *only*. Furthermore, there was a clear, gradual developmental trend to the processing of sentences with *only*. For sentences with pre-subject and pre-verbal *only*, the youngest age group of participants made a large number of errors, but this error rate decreased monotonically across age groups. An inspection of the means in Table 2 suggests that adults made as few errors for sentences with *only* as for ones without *only*. Whereas in Experiment 1 there was a higher error rate for pre-verbal than pre-subject *only*, no such differences were obtained in the present experiment. This may well be a consequence of differences between the two experimental tasks, and we will return to this point in Section 5.

Next we focussed on responses produced by adults and 6- to 7-year-olds in order to discriminate between our account of children's errors and that of Crain et al. We predicted that young children would make errors by neglecting contrast information, and judge Fig. 1A,C–E to be true for sentences with *only*. Crain et al. predicted that young children would make errors by misanalyzing the scope of *only*. Consequently, they should judge Fig. 1C rather than Fig. 1D to be true for sentences with pre-subject *only*, and judge Fig. 1D rather than Fig. 1C to be true for sentences with pre-verbal *only*. Thus, erroneously judging Fig. 1C or Fig. 1D to be true for sentences with *only* is consistent with both accounts. However, the accounts make different predictions concerning errors for Fig. 1E. Our account predicts a high error rate for this picture whereas Crain et al.'s account does not predict any errors for this picture. Table 4 shows the frequency with which 6- to 7-year-olds and adults judged individual pictures to be true for sentences with and without *only*.

Table 4
Mean percentage selection (with standard deviations) of each picture by adults and 6- to 7-year-olds when evaluating sentences without *only*, with pre-subject *only* and with pre-verbal *only* in Experiment 2

Age group	Sentence	Picture					
		A	B	C	D	E	F
Adults	Without *only*	100.0* (0.0)	0.0 (0.0)	94.2* (10.7)	93.3* (13.3)	93.3* (13.3)	0.0 (0.0)
	Pre-subject *only*	100.0* (0.0)	0.0 (0.0)	3.8 (9.2)	91.3* (14.0)	1.0 (4.9)	0.0 (0.0)
	Pre-verbal *only*	100.0* (0.0)	0.0 (0.0)	85.6* (24.7)	4.8 (10.0)	0.0 (0.0)	0.0 (0.0)
6–7 years	Without *only*	100.0* (0.0)	0.0 (0.0)	93.3* (24.0)	94.2* (24.0)	90.4* (27.5)	0.0 (0.0)
	Pre-subject *only*	100.0* (0.0)	1.0 (4.9)	66.3 (37.4)	77.9* (32.7)	48.1 (41.2)	0.0 (0.0)
	Pre-verbal *only*	100.0* (0.0)	1.0 (4.9)	87.5* (25.7)	50.0 (41.8)	42.3 (43.5)	0.0 (0.0)

Asterisks indicate pictures that match with sentence meaning.

Once again, both age groups produced a high frequency of correct responses and made few errors for sentences without *only*. Consistent with our predictions, 6- to 7-year-olds often judged Fig. 1A–E to be true for sentences with *only*. Furthermore, 6- to 7-year-olds judged Fig. 1E to be true on 48% of trials for pre-subject *only* and on 42% of trials for pre-verbal *only*. Such responses were consistent with our account, and unpredicted by Crain et al.'s account. In addition, there were 62.3% errors for Fig. 1C for pre-subject *only* and 42.3% errors for Fig. 1D for pre-verbal *only* that were consistent with both theoretical accounts. Adults made few errors by erroneously judging Fig. 1C, Fig. 1D or Fig. 1E to be true for sentences with *only*. Thus, it appears that children often made errors that were consistent with our account and inconsistent with the alternative Crain et al. account.

In order to conduct a more rigorous test of the hypotheses concerning the nature of children's comprehension errors we returned to the categorized data in Table 3 and compared the frequency with which adults and 6- to 7-year-olds produced *no scope* responses or made scope analysis errors for sentences with *only*. Data were analyzed using two 2 (age group) $\times$ 2 (sentence type) $\times$ 2 (response type) ANOVAs. Six- to 7-year olds made fewer errors than adults ($F_1(1, 50) = 59.84$, $P < 0.001$; $F_2(1, 22) = 147.24$, $P < 0.001$). There was no main effect of sentence type ($Fs < 1$), however, participants produced overall more *no scope* errors than scope analysis errors ($F_1(1, 50) = 13.70$, $P < 0.001$; $F_2(1, 22) = 22.97$, $P < 0.001$).

There was no interaction of age group and sentence type ($Fs < 1$), and the interaction of sentence and response type was not significant ($F_1 < 1$; $F_2(1, 22) = 2.16$, $P > 0.05$). However, there was a significant interaction of age group and response type ($F_1(1, 50) = 17.83$, $P < 0.001$; $F_2(1, 22) = 30.99$, $P < 0.001$). Post-hoc tests indicated that 6- to 7-year-olds produced more *no scope* than scope analysis errors ($P < 0.001$) but that adults did not differ in the production of *no scope* and scope analysis errors ($P > 0.05$). The three-way interaction of age group, sentence and response type was not significant ($F_1(1, 50) = 2.08$, $P > 0.05$; $F_2(1, 22) = 2.16$, $P > 0.05$). Thus, these results indicate that 6- to 7-year-olds principally made errors in which they neglected contrast information rather than errors in which they misanalyzed scope, with no difference in errors of these types for adults.

In summary, we replicated the key findings from Experiment 1. All participants produced a high frequency of correct responses for sentences without *only*, but young children made a large number of errors when evaluating sentences with pre-subject and pre-verbal *only*. The error rate decreased across age groups, with no evidence of a critical period for the acquisition of adult sentence processing strategies. An inspection of Table 2 suggests that adults made as few errors for sentences with *only* as they did for sentences without *only*. In contrast to Experiment 1, the error rates for sentences with pre-subject and pre-verbal *only* did not differ.

Participants made errors that were consistent with them misanalyzing scope or neglecting contrast information. However, 6- to 7-year-olds made a large number of errors that were wholly consistent with them failing to process contrast information, but were inconsistent with them misanalyzing scope. A formal analysis of error rates indicated that young children made significantly more errors in which they ignored contrast information than errors of scope analysis, with no such difference for adults.

4. Experiment 3

Experiments 1 and 2 clearly demonstrated that 6- to 7-year-olds made substantially more errors when evaluating sentences with *only* because they ignored contrast information rather than because they disregarded syntactic restrictions on scope. However, our experiments still differ in one important respect from the study conducted by Crain et al. Crain et al. examined the performance of children aged between 3 and 6 years old, whereas our youngest participants were aged between 6 and 7 years. As just over half the participants (i.e. 21 of 38 children) produced responses in which they misanalyzed the scope of sentences with pre-subject and pre-verbal *only*, it is possible that this effect was restricted to the youngest participants in the study. Therefore, we may obtain similar effects by examining the performance of a younger group of children than were examined in Experiments 1 and 2.

In Experiment 3, we examined the performance of 4- to 5-year-old pre-school children on a modified version of the forced-choice sentence–picture evaluation task used in Experiment 1. The task was modified in two important respects in order to make it practicable for this young age group. First, we reduced the number of trials in order to make the experiment shorter and easier for each child to complete. Second, sentences were re-phrased as questions (i.e. *Is the fireman holding a hose?*) to reduce the difficulty of the task. Thus, participants were asked questions about each picture rather than being asked to evaluate a statement with respect to each picture.

As in the previous experiments, we examined the pattern of correct responses for sentences with and without *only*, and compared the frequency of errors in which children misanalyzed sentences with *only* as having the same meaning as counterparts without *only* or else misanalyzed the scope of the focus particle.

4.1. Method

4.1.1. Participants
Twenty 4- to 5-year-olds from the Mickleover and Lonsdale Nursery Schools run by University of Derby, and Brisley Primary School participated in the study.

4.1.2. Materials and design
In order to reduce the length and difficulty of each test session, we selected six of the 12 sentences used as materials in Experiments 1 and 2. We used those sentences involving characters and actions that were most likely to be clearly understood by young children (e.g. characters such as firemen, policemen, nurses, and mothers, and actions such as carrying or holding objects). Again, in order to reduce the difficulty of the task, sentences were re-phrased as questions (e.g. Is the fireman holding a hose? Is only the fireman holding a hose?). The question materials used are listed in Appendix B.

Sentences were divided into three lists. Each list contained one version of each sentence and two instances of each sentence type. We used the force-choice judgement task. Participants were read each sentence in a fixed order and asked to indicate if the sentence was a true or false description of each of six corresponding pictures. The pictures were shown one at a time in a pseudo-random order. The independent variable was sentence

type: sentences without *only* and with either pre-subject or pre-verbal *only*. The dependent variable was the frequency of *no scope*, *NP scope*, *VP scope* and *other* responses.

4.1.3. Procedure

Participants were tested individually within the nursery. Each test session lasted 10–20 minutes. At the beginning of each trial, the experimenter described the characters and actions involved in the drawings, and then asked the child to name each character and object (the experimenter corrected the child if he or she made an error). This ensured that participants understood which characters and objects were referred to in sentences and knew the identity of the secondary characters and nature of contrasting objects. The experimenter began each trial by asking the child a question that was obviously false (e.g. Is the fireman holding a banana?) in order to familiarize the child with giving 'no' responses and to ensure that the child was attending to the task. The experimenter then asked the child to indicate if the question corresponding to the experimental sentence was true or false of each of the six pictures. A second experimenter recorded the child's responses.

4.2. Results and discussion

Responses were coded into *no scope*, *NP scope*, *VP scope* and *other* categories. The mean percentage and standard deviations for each type of response for each sentence are shown in Table 5.

Most responses were categorized as *no scope*, *NP scope* or *VP scope*. For 14.2% of trials participants made incorrect responses to Fig. 1A, Fig. 1B or Fig. 1F. For the others they made contradictory responses to pictures representing contrast information.

We first analyzed the frequency of correct responses for each sentence type using two one-way repeated measures ANOVAs, one treating subjects as the random variable (F_1), and the other treating stimuli items as the random variable (F_2). There was a significant effect of sentence type ($F_1(2, 38) = 22.08$, $P < 0.001$; $F_2(2, 10) = 22.45$, $P < 0.001$), with more correct responses for sentences without *only* than pre-subject *only* ($F_1(1, 19) = 12.67$, $P < 0.001$; $F_2(1, 10) = 7.31$, $P < 0.05$), or pre-verbal *only* ($F_1(1, 19) = 52.64$, $P < 0.001$; $F_2(1, 10) = 76.59$, $P < 0.001$). Finally, there were more correct responses for pre-subject *only* than for pre-verbal *only* ($F_1(1, 19) = 6.86$, $P < 0.05$; $F_2(1, 5) = 7.98$, $P < 0.05$).

The high rate of correct responses for sentences without *only* demonstrated that participants could understand and perform the task. However, children found sentences

Table 5
Mean percentage correct responses (with standard deviations) for 4- to 5-year-olds when evaluating sentences without *only*, with pre-subject *only* and with pre-verbal *only* in Experiment 3

Sentence	No scope	NP scope	VP scope	Other
Without *only* (no scope = correct response)	70.0 (34.0)	0.0 (0.0)	5.0 (15.4)	25.0 (34.4)
Pre-subject *only* (NP scope = correct response)	25.0 (34.4)	30.0 (41.0)	5.0 (15.4)	40.0 (34.8)
Pre-verbal *only* (VP scope = correct response)	37.5 (45.5)	12.5 (27.5)	5.0 (15.4)	45.0 (39.5)

Table 6
Mean percentage selection (with standard deviations) of each picture by 4- to 5-year-olds when evaluating sentences without *only*, with pre-subject *only* and with pre-verbal *only* in Experiment 3

Age group	Sentence	Picture					
		A	B	C	D	E	F
4–5 years	Without *only*	95.0* (15.4)	5.0 (22.4)	97.5* (11.1)	92.5* (18.3)	92.5* (18.3)	7.5 (24.5)
	Pre-subject *only*	87.5* (27.5)	10.0 (26.2)	57.5 (40.6)	80.0* (34.0)	52.5 (47.2)	5.0 (22.4)
	Pre-verbal *only*	82.5* (29.4)	12.5 (27.5)	60.0* (41.7)	75.0 (34.4)	57.5 (43.8)	5.0 (22.4)

Asterisks indicate pictures that match with sentence meaning.

with *only* more difficult to correctly evaluate than sentences without *only*. Note that as with Experiment 1 sentences with pre-verbal *only* produced a higher rate of errors than ones with pre-subject *only*.

We next examined the frequency with which individual pictures were judged true for each sentence. Mean frequency data are shown in Table 6.

The frequency data suggest that pre-schoolers had a strong understanding of the correct responses for sentences with and without *only*. For sentences without *only* they produced a high rate of correct responses and relatively few errors. They also produced a high frequency of responses in which they correctly judged Fig. 1A,D to be true for pre-subject *only*, and judged Fig. 1A,C to be true for pre-verbal *only*. Thus, it cannot be argued that pre-schoolers simply failed to understand the meaning of the focus particle in any context. Rather, the data suggest that they have a reasonably good understanding of the meaning of *only* but are prone to making interpretative errors. The critical question concerns whether these errors are made by misanalyzing scope or by failing to attend to contrast information.

As before we expected that if participants made errors by misanalyzing scope then they would erroneously judge Fig. 1C rather than Fig. 1D to be true for pre-subject *only* sentences and judge Fig. 1D rather than Fig. 1C to be true for pre-verbal *only* sentences. By contrast, if participants made errors by neglecting contrast information then we expected them to judge Fig. 1A,C–E to be true for sentences with *only*. The frequency with which participants selected Fig. 1E would discriminate between the alternative theoretical positions.

An inspection of Table 6 indicates that pre-schoolers often judged Fig. 1A,C–E to be true for sentences with *only*, as predicted by our account. Children made 57.5% errors by erroneously judging Fig. 1C to be true for sentences with pre-subject *only*, and 75% errors by erroneously judging Fig. 1D to be true for sentences with pre-verbal *only*, with these errors consistent with either theoretical position. Critically, they also judged Fig. 1E to be true for 52.5% of trials for pre-subject *only*, and for 57.5% of trials for pre-verbal *only*. Errors of this type were consistent with children having failed to attend to contrast information but offered no support to an alternative account in which they had misanalyzed scope.

We next conducted a formal test of pre-schoolers' relative tendency to make errors consistent with either failing to attend to contrast information or by misanalyzing scope by

returning to our categorization data in Table 6 and comparing the frequency of *no scope* and scope misanalysis errors for sentences with *only*. We employed two 2 (sentence type) × 2 (response type) repeated measures ANOVAs. The main effect of sentence type was not reliable ($F_1(1, 19) = 4.75, P < 0.05; F_2 < 1$), thus an analysis of categorized data did not support the observation from frequency data that more scope analysis errors were made for pre-verbal than pre-subject *only*. However, there was a significant main effect of response type ($F_1(1, 19) = 4.54, P < 0.05; F_2(1, 5) = 11.09, P < 0.05$). Participants made almost three times more errors in which they assigned a *no scope* analysis than errors in which they misanalyzed the scope of *only*. There was no interaction of sentence and response type ($Fs < 1$). These results indicated that pre-schoolers principally made errors by neglecting contrast information rather than errors in which they misanalyzed scope.

The results from this experiment replicated our findings for 6- to 7-year-olds in Experiments 1 and 2. Pre-schoolers made a high rate of correct responses for sentences without *only*. For sentences with *only*, the frequency data indicated that they often erroneously judged Fig. 1E to be true, with this response exclusively consistent with an account in which pre-schoolers neglected contrast information and inconsistent with one in which they misanalyzed scope. A formal analysis of categorized data supported this conclusion. Pre-schoolers made significantly more errors by neglecting contrast information than by misanalyzing scope. Thus, our results suggest that pre-schoolers predominately made errors because they fail to instantiate explicit contrast information in their discourse model.

5. General discussion

Our aim in conducting these studies was to determine whether adults and children interpreted sentences with *only* by constructing discourse models that necessarily included focus and contrast sets, and whether they used syntactic cues to restrict the scope of the focus particle. There were two main findings. First, children and adults had more difficulty in evaluating sentences with than without *only*. All participants found it easy to correctly evaluate sentences without *only*. In Experiment 1 young children made a large number of errors for pre-subject *only*, and the error rate decreased across ages to the extent that adults made as few errors for sentences with pre-subject *only* as they did for ones without *only*. Experiment 1 results for pre-verbal *only* were unexpected, however, as all participants made a uniformly high rate of errors. In Experiment 2 young children also made a large number of errors for sentences with *only*, and again the error rate decreased across ages with adults making as few errors for sentences with *only* as they did for ones without *only*. Finally, in Experiment 3 pre-schoolers made a large number of errors for pre-subject *only* and even more for pre-verbal *only*.

Before considering the significance of these findings we must rule out one theoretically uninteresting alternative explanation of our results. It could be argued that we obtained different results for children and adults because children had difficulty in understanding sentences and pictures or generally found it difficult to perform the experimental task. However, our data do not support this view. We established that all ages produced a high rate of correct responses for sentences without *only*. This indicated that they understood

and could perform the experimental task. Participants experienced no difficulty for sentences without *only* but made errors when *only* indicated contrastive focus, and results for Experiments 1 and 3 suggest that participants (of all ages) had greater difficulty evaluating sentences with pre-verbal than pre-subject *only*. Thus, the results demonstrate that young children experienced difficulty when processing sentences containing focus particles that required the reader to form contrastive focus, and that the level of difficulty depended on the syntactic position of the particle. Adults also experienced difficulty when instantiating contrast sets under certain conditions.

The second key finding related to the types of error that were made. Crain et al. (1994) claimed that young children disregard syntactic restrictions on scope when processing sentences with *only*, and adopt the same analysis regardless of the surface position of the particle. Following this account we would expect young children predominantly to make errors in which they misanalyzed scope. That is to say, Crain et al. would expect children to assign the same meaning to sentences in which *only* occurs in different syntactic positions. We advocated an alternative account and argued that children principally make errors when evaluating sentences with *only* because they do not include explicit sets of alternatives within their discourse model. Thus, we predicted that children often would make errors by neglecting contrast information and misanalyze sentences with *only* as having the same meaning as counterparts without *only*.

Participants made errors in which they misanalyzed scope and, in addition, errors indicating that they had failed to include explicit sets of alternatives within their discourse model. Crucially, in all three experiments there was a high frequency of errors for the youngest children (4- to 5-year-olds in Experiment 3 and 6- to 7-year olds in Experiments 1 and 2) that unambiguously could be attributed to them neglecting contrast information. When we categorized children's erroneous responses as either ones in which they misanalyzed scope or ones in which they neglected contrast information, we found that errors of neglecting contrast information predominated. Crain et al.'s theory predicted that young children predominantly would make errors by misanalyzing scope. As this does not appear to be the case, Crain et al.'s theory cannot provide a general explanation of why children have difficulty in processing sentences with focus particles. Rather, the results support our view that children often make errors when processing sentences with focus particles because they fail to mentally represent contrast information. Note also that it cannot simply be the case that children lack any understanding of the meaning of *only* and ignore it when processing sentences as the frequency data show that on many trials they correctly evaluated the match between sentences with *only* and pictures depicting contrast information.

Our results differed from those obtained in previous studies (e.g. Crain et al., 1992; Philip & Lynch, 1999). In particular, Crain et al. claimed to demonstrate that children misanalyzed sentences with pre-subject *only* as having the same meaning as counterparts with pre-verbal *only*. Note that Experiment 3 ruled out the possibility that differences between the current findings and those of Crain et al. are due to differences in the age of the participant population tested. In Experiment 3 we obtained a similar pattern of results for 4- and 5-year-olds as found for 6- and 7-year-olds in Experiment 1.

Note also that in Experiments 1 and 3 we employed a different task from that used by Crain et al. Our participants compared the target sentence with each of six pictures in turn

and made forced-choice decisions about whether the sentence and picture matched. In contrast Crain et al.'s participants simultaneously viewed all of the alternative pictures before selecting the one that was consistent with the test sentence. One possibility was that viewing the alternative pictures in parallel provided a strong cue to using contrast information. This cue may have been absent in the forced-choice version of the task, inflating the incidence of errors in which participants failed to take account of contrast information. Therefore, in Experiment 2 we used a task in which participants viewed all six pictures simultaneously and identified which ones were consistent with the target sentence. We replicated the principle findings from Experiment 1, ruling out the possibility that differences between our results and Crain et al.'s were due to employing a particular experimental task.

The differences between our results and those obtained by Crain et al. are most likely due to differences in experimental design. Critically, our experiments differed from previous studies by including control conditions that enabled us to discriminate between errors of scope analysis and errors in which participants neglected contrast information. Crain et al.'s study lacked these control conditions. They compared sentences with pre-subject *only* and pre-verbal *only* against pictures that were consistent with direct object scope and inconsistent with subject noun-phrase scope. When participants judged that sentences with pre-subject *only* matched with the picture, Crain et al. interpreted this as evidence that they had misanalyzed scope. However, a counterpart sentence without *only* also would be judged to be true of the depicted events. Thus, participants may not have misanalyzed scope, but instead may have evaluated sentences without taking account of contrast information, and analyzed sentences with *only* as sharing the meaning of counterparts without *only*. This alternative explanation is consistent with our present findings.

There is one aspect of our results that we have still to consider. The results for pre-verbal *only* differed across experiment tasks. In Experiment 1, all ages made a large number of errors for pre-verbal *only* and in Experiment 3 pre-schoolers made more errors for pre-verbal *only* than for pre-subject *only*. In contrast, in Experiment 2 participants produced a similar number of errors for pre-verbal and pre-subject *only*. Clearly, sentences with pre-verbal *only* were more difficult to evaluate than ones with pre-subject *only* and this difficulty was acerbated by the forced-choice task.

Participants evidently found it easier to evaluate contrastive focus on the free-choice rather than forced-choice task. We had considered it likely that simultaneously viewing all of the alternative pictures in the free-choice task would provide a strong cue to using contrast information. However, this was not supported by the data. Had the free-choice task cued participants to using contrast information then we might have expected this to reduce the frequency of responses in which they analyzed sentences with *only* as having the same meaning as counterparts without *only*. However, this was not the case. One possible alternative explanation is that the forced-choice task caused participants to make a series of independent comparisons of the target sentence with each picture, without tracking the contingencies between decisions. In contrast, the free-choice task required them to make a single set of comparisons in which the contingencies between decisions are obvious. That is to say, when all the alternatives are simultaneously available the selection of one picture necessarily rules out the selection of others.

We can think of two non-mutually exclusive explanations for why pre-verbal *only* produced more errors than pre-subject *only*. First, participants may have found pre-verbal *only* difficult to evaluate because, even with scope restricted to the verb-phrase, the sentence is ambiguous between alternative scope analyses. Alternatively, participants may have experienced more difficulty in processing contrast information for pre-verbal *only*. Pre-subject *only* indicated a contrast between the person described in the sentence and some alternatives, whereas pre-verbal *only* indicated a contrast between an event described by the sentence and some alternatives. Participants may have found it easier to evaluate a contrast between two characters than a contrast between two events. At present we cannot discriminate between these two possibilities.

This leads onto a more general consideration of how contrast information is mentally represented and why young children may not always instantiate explicit contrast information in their discourse model. We noted in Section 1 that when the contrast set is not made explicit the reader or listener must infer one using pragmatic knowledge or knowledge about the prior referential context (Frazier, 1999; Sedivy, 2002). In the absence of disambiguating referential context how a reader or listener mentally represents contrastive focus may depend on the ease with which they can infer a set of alternatives. This will depend in part on the knowledge base of the reader or listener, and on semantic and pragmatic properties of the focus element. For example, Clifton et al. (2000) examined syntactic processing of sentences with *only* when the focus element was part of a high or low contrast pair. In sentence (9) *only* takes scope over a subject noun-phrase (*the winners*) that has a high contrast counterpart in *the losers*, whereas in (10) it takes scope over a subject noun-phrase (*the pirates*) without a high contrast counterpart.

9. Only the winners offered some money went on television that evening.
10. Only the pirates offered some money surrendered their ship.

Clifton et al. argued that readers should find it easier to infer alternatives for a high contrast element than for a low contrast element. That is, the set of winners will have an obvious contrast set whereas the set of pirates will not. Although Clifton et al. found that contrastive focus did not modulate parsing decisions for syntactically ambiguous sentences, high and low contrast pairs did influence overall sentences reading times. It seems likely that the ease with which a contrast set can be inferred will determine how readily it is instantiated as part of the discourse model.

Our explanation of the current findings is as follows. Readers and listeners may not always instantiate explicit sets of alternatives in the discourse model for sentences with focus particles. They may find it easy to include a contrast set when one is available in the prior referential context or can easily be inferred. However, if an explicit contrast set is not available or cannot easily be inferred, then they may either fail to instantiate any contrast set in their discourse model or they may include an underspecified set of alternatives. Whichever is the case, the failure to instantiate an explicit set of alternatives will impair performance on sentence–picture comparison tasks. Participants who fail to mentally represent an explicit contrast set may make errors in which they process a sentence with a focus particle as if it had the same meaning as a counterpart without a focus particle. In our experiments, children made more errors of this nature than did adults. Thus, the results are

consistent with an account in which children find it difficult to mentally represent contrast sets as part of their discourse model. Importantly, we are not arguing that children never include contrast sets or that such representations are exclusive to children. In fact, we imagine that there will be occasions when adults construct discourse models that do not include explicit contrast sets. Underspecified representations simply may predominate for young children because they have less well-developed pragmatic knowledge than older children or adults. Because of restricted knowledge they have more difficulty in inferring a set of alternatives and therefore are more likely to fail to include one in their discourse model. To conclude, we propose that children and adults differ in their ability to infer contrast sets and this affects how they construct discourse models.

Our approach to focus particles has much in common with the approach taken by Brooks and Braine (1996) and Geurts (2001) to the acquisition of universal quantifiers. Like them we have outlined an account of children's linguistic errors that does not require a discontinuity between child and adult language processing, or that children lack specific syntactic rules. Rather, errors occur as a consequence of mapping linguistic form onto a semantic representation (discourse model). Children often construct representations that include only information that is made explicit in the sentence, whereas adults often construct representations that include implied information.

Consider the Brooks and Braine (1996) account. They proposed that children make errors when evaluating sentences with *each* because they at first analyze such sentences as mapping onto an exhaustive and distributive representation rather than the non-exhaustive distributive representation preferred by adults. Thus, on encountering a sentence like *Each of the boys was carrying a box*, children construct a discourse model with a one-to-one correspondence between boys and boxes and no boxes leftover rather than a one-to-one correspondence that permits leftover boxes. Importantly, the construction of a non-exhaustive model requires the inclusion of information that is not made explicit (i.e. information about extra entities).

Geurts (2001) proposed that children make errors because they at first assign a weak reading to all quantifiers, whereas universal quantifiers take a strong reading. The strong reading for a sentence like *Most of the men carried a sofa* should include representations of the set of many men that carried a sofa and are made explicit in the sentence, and an implied set of fewer men that did not carry a sofa. A weak reading will not include the implied set of men who did not carry a sofa. Geurts argued that children make errors on comprehension tests because of the nature of the underspecified semantic representation they construct.

More tentatively, we see links with Noveck's recent accounts of performance differences between children and adults on logical reasoning tasks (Noveck, 2001). Noveck found that adults took statements including *some* to imply that "not all" was the case, but 7-year-olds adopted a logical reading of the quantifier in which statements such as *Some giraffes have long necks* are compatible with ones like *All giraffes have long necks*. Noveck explained this in terms of the development of pragmatic inference, and his account is compatible with the present results. Just as Noveck found that children fail to treat *some* as implying *not all*, we have identified cases in which children fail to treat *only x are y* as implying that *not-x are not y*.

Finally, we can relate our findings to broader considerations of children's processing of scope ambiguities. We outlined in Section 1 the Crain et al. (1994) account of how

children and adults process scope ambiguities created by focus particles. This was a detailed account that made testable predictions but implied a discontinuity in language development. Adults were claimed to process ambiguities following the Referential theory, whereas children were claimed to process sentences following principles that support the acquisition of alternative scope analyses. With the present studies we have produced clear evidence against one aspect of their account – that children do not use knowledge of syntactic structure to restrict scope.

Recent studies by Musolino and colleagues (Lidz & Musolino, 2002; also Musolino, in press; Musolino et al., 2000) support our view that children use syntactic cues to restrict scope and point to an alternative account of how they might deal with ambiguities. Lidz and Musolino examined how child and adult speakers of English and Kannada (a language from Southwest India) understood sentences containing numeric quantifiers and negation, e.g. *The policeman didn't find two guys*. This sentence can be interpreted with either the quantifier (*two guys*) taking scope over the negation (*not*), or with negation taking scope over the quantifier. With the quantifier taking scope over the negation (i.e. taking wide scope) it asserts that the policeman failed to find two specific persons. By contrast, with negation taking wide scope it asserts that the men found by the policeman did not number two (because, for instance, the policeman only found one man, or because he found three).

Lidz and Musolino found that adults judged sentences to be acceptable in contexts favouring either a wide-scope negation or wide-scope quantifier analysis, suggesting that both scope analyses were available to them, and that they selected the one that matched with context. However, children only judged sentences to be acceptable in contexts favouring wide-scope negation, suggesting that they computed just this analysis of the sentence. Lidz and Musolino argued that English-speaking children either used a linear order strategy to assign wide scope to the leftmost operator (e.g. Fodor, 1982; Johnson-Laird, 1969), or syntactic knowledge to assign wide scope to the particle occupying the dominant position in the parse tree (e.g. Ioup, 1975; Kurtzman & MacDonald, 1993). Word order differences between English and Kannada enabled them to rule out the linear order explanation and conclude that children used syntactic structure to determine scope relations. They could not account for children and adults showing different knowledge of scope relations, but speculated that it may be due to children lacking linguistic knowledge of the unavailable scope analysis (e.g. Kramer, 2000), or processing limitations preventing them from considering alternatives to their default scope analysis (e.g. Frazier, 1999; Reinhart, 1995; Trueswell, Sekerina, Hill, & Logrip, 1999).

Clearly further work is needed to determine whether there are qualitative differences in how children and adults interpret ambiguous sentences. This would contrast approaches that emphasize the role of referential context in guiding sentence interpretation (Crain et al., 1994; Sedivy, 2002; Sedivy, Tannenhaus, & Carlson, 1998) and alternative approaches that emphasize structural principles (Frazier, 1999; Ioup, 1975). First, there is a need to establish whether analyses are considered in serial or in parallel. Following Crain et al.'s application of the Referential theory to scope processing, we would expect adults to consider alternative analyses in parallel before opting for the one that matches with the prior referential context or requires fewest referential presuppositions. Alternative parallel processing accounts include Race-based accounts (e.g. Van Gompel, Pickering, & Traxler, 2000) and Constraint-satisfaction accounts (e.g. Sedivy, 2002; Sedivy et al., 1998).

By contrast serial processing accounts (e.g. Frazier, 1999) require that adults preferentially adopt one of the possible analyses and reanalyze if it proves to be incorrect. At present the data do not discriminate between serial and parallel processing accounts.

To conclude, the present experiments demonstrated that children and adults are sensitive to restrictions on scope imposed by structural factors, and that children often fail to instantiate explicit contrast sets. Furthermore, we have demonstrated that children often do not form adult-like interpretations of sentences with focus particles principally because they fail to employ pragmatic information to infer explicit contrast sets, rather than because they fail to employ knowledge of syntactic structure.

Acknowledgements

Appendix A. Sentence materials used in Experiment 1 and 2

(Only) the woman is (only) walking a dog.
(Only) the boy is (only) flying a kite.
(Only) the boy is (only) caught a fish.
(Only) the man is (only) carrying a briefcase.
(Only) the fireman is (only) holding a hose.
(Only) the lollipop lady is (only) stopping a car.
(Only) the boy is (only) playing with a ball.
(Only) the nurse is (only) lifting a bucket.
(Only) the girl is (only) stroking a horse.
(Only) the woman is (only) pushing a pram.
(Only) the postman is (only) delivering a letter.
(Only) the dog is (only) chasing a cat.

Appendix B. Sentence materials structured as questions used in Experiment 3

Is (only) the fireman (only) holding a hose?
Has (only) the boy (only) caught a fish?
Is (only) the boy (only) flying a kite?
Is (only) the man (only) carrying a bag?
Is (only) the woman (only) pushing a pram?
Is (only) the nurse (only) carrying a bucket?

References

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, *75*, 87–106.

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

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

Bucci, W. (1978). The interpretation of universal affirmative propositions: a developmental study. *Cognition*, *6*, 55–77.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (Ed.), *Subject and topic*. New York: Academic Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behaviour*, *12*, 335–359.

Clark, H. H., & Clark, E. V. (1973). *Psychology and language*. New York: Harcourt, Brace, Jonanovich.

Clifton, C., Jr., Bock, J., & Rado, J. (2000). Effects of the focus particle *only* and intrinsic contrast on comprehension of reduced relative clauses. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process*. Amsterdam: Elsevier.

Crain, S. (2000). Sense and sense ability in child language. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University conference on language development*, Somerville, MA: Cascadilla Press.

Crain, S., Ni, W., & Conway, L. (1994). Learning, parsing and modularity. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum.

Crain, S., Philip, W., Drozd, K., Roeper, T., & Matsuoka, K (1992). *Only in child language*. Unpublished manuscript, University of Connecticut, Storrs, CT.

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. Kartunnen, & A. M. Zwicky (Eds.), *Natural language parsing: psychological, computational, and theoretical perspectives*. Cambridge: Cambridge University Press.

Crain, S., & Thornton, R. (1998). *Investigations in universal grammar: a guide to research on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

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

Donaldson, M., & Lloyd, P. (1974). Sentences and situations: children's judgments of match and mismatch. In F. Bresson (Ed.), *Problemes actuels en psycholinguistique*. Paris: Centre National de Recherche Scientifique.

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

Drozd, K., & van Loosbroek, E (1998). *Dutch children's interpretations of focus particle constructions*. Poster presented at the 23rd annual Boston University Conference on Language Development, Boston, MA.

Fodor, J. D. (1982). The mental representation of quantifiers. In S. Peters, & E. Saarinen (Eds.), *Processes, beliefs, and questions*. Dordrecht: Reidel.

Frazier, L. (1999). *On sentence interpretation*. Dordrecht: Reidel.

Frazier, L., Pacht, J. M., & Rayner, K. (1999). Taking on semantic commitments II: collective versus distributive readings. *Cognition*, *70*, 87–104.

Freeman, N. H., & Stedmon, J. A. (1986). How children deal with natural language quantification. In I. Kurcz, G. W. Shugar, & J. H. Danks (Eds.), *Knowledge and language*. Amsterdam: Elsevier.

Garrod, S. C., & Sanford, A. J. (1994). Resolving sentences in a discourse context: how discourse representation affects language understanding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics*. New York: Wiley.

Geurts, B (2001). *Quantifying kids*. Paper presented at the Workshop on Acquisition Perspectives on Interfaces, University of Utrecht.

Halbert, A. M., Crain, S., Shankweiler, D., & Woodams, E (1995). *Interpretive use of emphatic stress by children*. Poster presented at the 8th annual CUNY Conference on Human Sentence Processing, Tucson, AZ.

Hanlon, C. C. (1986). Acquisition of set-relational quantifiers in early childhood. *Genetic, Social, and General Psychology Monographs*, *113*, 213–264.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Duxbury.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child*. London: Routledge.

Ioup, G. (1975). Some universals for quantifier scope. In J. Kimball (Ed.), (*4*). *Syntax and semantics*, New York: Academic Press.

Jackendoff, R. S. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Johnson-Laird, P. N. (1969). On understanding logically complex sentences. *Quarterly Journal of Experimental Psychology*, *21*, 1–13.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: Cambridge University Press.

Konig, E. (1991). *The meaning of focus particles*. London: Routledge.

Kramer, I. (2000). Interpreting indefinites: an experimental study of children's language comprehension. Doctoral dissertation, MPI Series in Psycholinguistics.

Kurtzman, H., & MacDonald, M. (1993). Resolution of quantifier scope ambiguities. *Cognition*, *48*, 243–279.

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

Link, G. (1983). The logical analysis of plurals and mass terms: a lattice-theoretical approach. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use and the interpretation of language*. Berlin: Walter de Gruyter.

Liversedge, S. P., Paterson, K. B., & Clayes, E. (2002). The influence of *only* on syntactic processing of "long" relative clause sentences. *Quarterly Journal of Experimental Psychology*, *55A*, 225–240.

Micham, D. L., Catlin, J., Van Derveer, N. J., & Loveland, K. A. (1980). Lexical and structural cues to quantifier scope relations. *Journal of Psycholinguistic Research*, *9*, 367–377.

Moxey, L. M., & Sanford, A. J. (1987). Quantifiers and focus. *Journal of Semantics*, *5*, 189–206.

Musolino, J (in press). Structure and Meaning in the acquisition of Scope. In L. Frazier, T. Roeper, & K. Wexler (Eds.), *Studies in Theoretical Psycholinguistics: Acquisition meets Semantics*. Dordrecht: Kluwer Academic Press.

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

Neimark, E. D., & Chapman, R. H. (1975). Development of the comprehension of logical quantifiers. In R. J. Falmagne (Ed.), *Reasoning: representation and process in children and adults*. Hillsdale, NJ: Erlbaum.

Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden paths: the contribution of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes*, *11*, 283–334.

Noveck, I. (2001). When children are more logical than adults: experimental investigations of scalar implicature. *Cognition*, *78*, 165–188.

Paterson, K. B., Liversedge, S. P., & Underwood, G. (1999). The influence of focus operators on parsing of *short* reduced relative clause sentences. *Quarterly Journal of Experimental Psychology*, *52A*, 717–738.

Paterson, K. B., Sanford, A. J., Moxey, L. M., & Dawydiak, E. (1998). Quantifier polarity and referential focus during reading. *Journal of Memory and Language*, *39*, 290–306.

Philip, W. (1995). *Event quantification in the acquisition of universal quantification*. Unpublished PhD thesis, University of Massachusetts, Amherst.

Philip, W., & Lynch, E. (1999). Felicity, relevance, and acquisition of the grammar of *every* and *only*. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University conference on language development*, Somerville, MA: Cascadilla Press.

Philip, W., & Takahashi, M. (1991). Quantifier spreading in the acquisition of *every*. In T. L. Maxfield, & B. Plunkett (Eds.), *University of Massachusetts occasional papers: papers in the acquisition of WH*. Amherst, MA: GLSA.

Reinhart, T. (1983). *Anaphora and semantic interpretation*. London: Routledge.

Reinhart, T. (1995). Interface strategies (OTS working papers in Linguistics), Utrecht University.

Reinhart, T. (1999). The processing cost of reference-set computation: guess patterns in acquisition (OTS working papers in Linguistics), Utrecht University.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, *1*, 75–116.

Rooth, M. (1996). Focus. In S. Lappin (Ed.), *The handbook of contemporary semantic theory*. Oxford: Blackwell.

Sanford, A. J., Moxey, L. M., & Paterson, K. B. (1996). Attentional focusing with quantifiers in production and comprehension. *Memory and Cognition*, *24*, 144–155.

Sedivy, J. C. (2002). Invoking discourse-based contrast sets and resolving syntactic ambiguities. *Journal of Memory and Language, 46*, 341–370.

Sedivy, J., Tannenhaus, M., & Carlson, G. (1998). Using contrast sets to resolve reference on-line. *University of Massachusetts occasional papers in linguistics 21: proceedings of the workshop on focus*. Amherst, MA: GLSA.

Trueswell, J. C., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition, 73*, 89–134.

Van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2000). Unrestricted race: a new model of syntactic ambiguity resolution. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process*. Oxford: Elsevier.