The kindergarten-path effect revisited: children's use of context in processing structural ambiguities

WEIGHALL, A. R.

Published version

WEIGHALL, A. R. (2007). The kindergarten-path effect revisited: children's use of context in processing structural ambiguities. Journal of experimental child psychology, 99, p. 75.

JECP 07-00094

Running head: CHILDREN'S USE OF CONTEXT IN PROCESSING STRUCTURAL AMBIGUITIES

The kindergarten-path effect revisited: Children's use of context in processing structural ambiguities.

Anna R. Weighall

Psychology, Sheffield Hallam University

Send correspondence to:

Dr. Anna R. Weighall

Psychology

Faculty of Development and Society

Sheffield Hallam University

Collegiate Crescent Campus

Sheffield S10 2BP, UK

a.r.weighall@shu.ac.uk

Tel: + 44 114 225 5563

Fax: + 44 114 225 2430

Please cite as: Weighall, A.R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. Journal of Experimental Child Psychology, 99 (2008) 75–95

Abstract

Research with adults has shown that ambiguous spoken sentences are efficiently resolved, exploiting multiple cues – including referential context – in order to select the intended meaning (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). Paradoxically, children appear to be insensitive to referential cues when resolving ambiguous sentences, relying instead upon statistical properties intrinsic to the language such as verb biases (Trueswell, Sekerina, Hill & Logrip, 1999). The possibility that children's insensitivity to referential context may be an artifact of the experimental design used in previous work is explored with 60 children aged 4- to 11-years-old. An act-out task was designed in order to discourage children from making incorrect pragmatic inferences, and to prevent premature and ballistic responses by enforcing delayed actions. Performance on this task was directly compared with the standard act-out task used in previous studies. The results suggest that young children (5-year-olds) do not utilize contextual information even under conditions designed to maximize their use of such cues, but that adult-like processing is evident by around the age of 8-years-old. These results support and extend previous findings by Trueswell et al., (1999) and are consistent with a constraint-based learning account of children's linguistic development.

Keywords: syntactic ambiguity resolution; act-out task; language development, spoken language comprehension; prepositional phrases

Introduction

Spoken language is often ambiguous with some syntactic structures proposing a myriad of alternative interpretations. When faced with a sentence like (1) the ambiguity is apparent with amusing consequences.

(1) Long lost sisters reunited after 18 years in checkout queue

In this example the possibility that the sisters were in a checkout queue for 18 years is entertaining but implausible (example taken from Altmann, 1998). Many syntactic ambiguities are less obvious, yet the adult sentence processing system resolves them seamlessly, and often without our conscious awareness that an ambiguity was ever present (see Altmann, 1998 for a comprehensive review of this literature). Adults make rapid use of multiple sources of information (e.g., discourse and visual context, frequency of co-occurrence, syntax, semantics, prosody) to disambiguate language (Eberhard, Spivey-Knowlton, Sedivy & Tanenhaus, 1995; Macdonald, Pearlmutter & Seidenberg, 1994; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). In contrast, children's ability to use such information has been disputed. Specifically, it has been suggested that young children do not use referential context in an adult-like way when faced with spoken syntactic ambiguity, and they must gradually learn which constraints to rely upon (Trueswell, Sekerina, Hill & Logrip, 1999). This paper systematically examines the extent to which children use referential context to disambiguate spoken sentences containing prepositional phrases ambiguities.

Several different classes of model have been proposed to explain how adults' sentence processing machinery (the parser) resolves syntactic ambiguity. These accounts agree that we rely upon some strategies or constraints to interpret the

language efficiently, given that ambiguity is resolved faster than would ever be possible if we waited until we had heard the whole sentence before committing to a representation.

Early *syntax-first* models propose that syntactic information is of primary importance during the early stages of processing, with extra-linguistic information (such as context) being considered only during the later stages once a misanalysis has taken place (Frazier & Fodor, 1978; see Frazier, 1987 for a review). In contrast, more recent work has favored a flexible and interactive *constraint-based* model of ambiguity resolution whereby different sources of information compete until the representation that has the best fit with all the available information is accepted and processed (Macdonald et al., 1994). The predictions made by these models have been investigated using prepositional phrase ambiguities such as (2).

(2) Put the frog on the napkin* into the box

(3) Put the frog that's on the napkin into the box

The first prepositional phrase (*on the napkin*) is temporarily ambiguous as to whether it conveys the destination of *the frog* (where the frog is to be moved to), or additional modifying information about it (it is currently on the napkin). In contrast sentence (3) is equally complex but unambiguous. The complementizer *that's* clearly indicates that *on the napkin* tells us where the frog is currently situated. Syntax-first models suggest that listeners will initially misinterpret the ambiguous first prepositional phrase *on the napkin* in *(2)* as conveying the destination of the frog on the basis of a parsing strategy called minimal attachment, a heuristic that states that the parser should prefer the syntactically simplest possible representation. Adherence to this heuristic leads to a

garden path effect whereby the listener or reader is led down the garden path to an incorrect interpretation which must be revised upon hearing *in the box.* There is evidence of such effects from adult reading studies which typically present single sentences in isolation (Britt, 1994; Ferreira & Clifton, 1986).

However, adults can avoid being led down the garden path when language is situated in a pragmatically appropriate context (Altmann, Garnham & Henstra, 1992; Tanenhaus et al., 1995). The *principle of referential success,* proposed by Altmann and Steedman (1988), predicts that sentences like (2) would result in a temporary ambiguity at the point indicated with an asterisk if there was only one frog present in the co-occurring context (e.g., preceding text, discourse, or the visual scene). In this case a modifier is not required in order to identify the intended referent as it would be pragmatically appropriate to say just *the frog.* Hence, the parser is led to misinterpret *on the napkin* as a destination – a commitment that must be revised upon encountering the second prepositional phrase. However, if two frogs are present a modifier is required in order to successfully establish reference - that is, *which* frog is to be moved. In this case the referential theory (and later constraint based models) predict that the ambiguity effect would be eliminated (or at least reduced). Essentially the principle of referential success states that the parser will prefer syntactic representations that successfully refer to distinct entities in the discourse.

This view has been supported by studies with adults (using spoken language) which demonstrate that visual context can prevent the initial destination interpretation of the first prepositional phrase (Spivey, Tanenhaus, Eberhard & Sedivy, 2002; Tanenhaus et al., 1995; but see Britt, 1994, and Ferreira & Clifton, 1986, for contrasting evidence with written language). Tanenhaus et al. (1995) monitored listeners' eye movements to objects in a visual array as spoken prepositional-phrase

sentences unfolded in real time. When there was just one potential referent participants fixated upon the incorrect destination (an empty napkin) on 55% of trials, shortly after hearing the first prepositional phrase suggesting that they initially considered a destination interpretation of the first prepositional phrase. In contrast they fixated the incorrect destination only 17% of the time when there were two referents. Importantly this did not differ significantly from the percentage of fixations on the incorrect destination in an unambiguous control sentence such as (3) suggesting that the presence of two referents eliminated the ambiguity effect. Contrary to the predictions made by syntax-first accounts this study clearly demonstrates that adults can integrate the context incrementally with the linguistic input to resolve syntactic ambiguities. This suggests that the adult parser operates in an interactive way that adheres to the referential principle, and exploits multiple cues to resolve syntactic ambiguity in an immediate and automatic fashion. Children's ambiguity resolution may not proceed under the same constraints as adults. In an influential study Trueswell at al. (1999) compared adults' and children's interpretations of sentences like (2) and (3). Closely following Tanenhaus et al. (1995) children's eye movements were monitored during listening whilst the order in which they acted out such sentences using an array of toy props was also recorded. Examples of the types of visual arrays and sentences used can be found in Figure 1 (pictures 1 and 2).

Trueswell et al. (1999) found that 5-year-old children behaved in a way that suggested they did not utilize the referential principle in an adult like manner. This finding poses interesting questions about the way in which the developing parser must learn to attend to relevant sources of information, and about the time-course of such development. In particular, children's eye movements revealed that they interpreted

the napkin as the destination most of the time. This was borne out in the children's actions, with children moving a frog to the empty napkin on over 60% of trials, irrespective of the number of referents. Thus, the first prepositional phrase was rarely interpreted as a modifier, and the presence of a supportive visual scene did not ameliorate children's difficulties with this ambiguity. This is in contrast to the adults who, like those studied by Tanenhaus et al. (1995), rarely considered a destination interpretation in the two-referent condition (as evidenced by the eye movement record) and performed the required actions almost flawlessly in all conditions. Furthermore, children's performance on unambiguous control sentences (*Put the frog that's on the napkin in the box*) was close to ceiling suggesting that the findings reflected the inability to resolve the ambiguity, rather than just the complexity of the sentences. Trueswell et al. (1999) dubbed this finding the *kindergarten-path effect*. They also tested a sample of 8-year-olds and although this group was quite heterogeneous it seemed that by this age performance was substantially adult-like.

These findings could be interpreted as evidence that children process language according to general parsing preferences, for example choosing the syntactically simplest representation (in line with *syntax-first* models e.g., Frazier, 1987; 1989), with subsequent re-analysis becoming more efficient as cognitive capacity increases with age. This position has been suggested by some to explain children's difficulties with other complex (but unambiguous) structures, such as relative clauses (e.g., Goodluck & Tavakolian, 1982; Goodluck, 1990; Tavakolian, 1981) but cannot account for research that has suggested that children are highly sensitive to the referential principle under some circumstances (Hamburger & Crain, 1982; Kidd & Bavin, 2002; Weighall & Altmann, 2001).

Alternatively, Trueswell et al. (1999) suggested that children may be unable to revise initial parsing commitments because of their more limited processing resources (compared to adults). More recently they have revised this suggestion to implicate a role for executive function in terms of the requirement to select between (and inhibit) competing representations (Trueswell, Papafragou & Choi, in press; Novick, Trueswell, & Thompson-Schill, 2005). Furthermore Trueswell et al. (1999) propose that children's difficulty with the type of constructions described here may lie in the probabilistic properties associated with the verb *put*. It is not permissible to say simply *Put the frog…*; a destination is required. Corpus data confirms that this verb is overwhelmingly likely to be combined with *in* or *on* and linked as a verb argument to the destination of the moved object (Trueswell et al. 1999). This means that upon hearing *put* it is highly probable that an intended destination will follow. Children may exploit this reliable statistical information to predict the upcoming language at the expense of the less consistent contextual information (e.g., number of potential referents); this would lead them to make the observed destination interpretation errors. As language users become more experienced contextual cues become a more salient and reliable source of disambiguating information, only then is the information used in initial syntactic representation. This explanation is consistent with constraint-based models of parsing, which see multiple cues as competing in ambiguity resolution (MacDonald etc al, 1994), and forms the basis of Trueswell and Gleitman's (2004) theory of constraint based language learning. Trueswell and colleagues characterize the developing parser as gradually learning the relevant constraints to attend to. The difference between adults and children arises because cues are differentially prioritized. Even in adults lexical biases are not always completely eliminated by

context as evidenced by the small number of errors occasionally made by adults in two referent conditions (Trueswell et al., 1999).

Further work by Snedeker and Trueswell (2004) confirmed the profound effect of verb biases upon children's ambiguity resolution by demonstrating that, when the lexical properties of the language are manipulated, children's biases can be switched to a modifier bias. When presented with a verb that is likely to be followed by an instrument, like *tickle* in (3), five year old children were very likely to arrive at an instrument-based interpretation of the sentence (that the fan was to be used to do the tickling). However, when faced with a verb that usually requires a modifier (e.g., *choose* in sentence 4) the pattern was reversed and they preferred a modifier interpretation (that the stick uniquely identifies a cow). In this respect children demonstrated an accurate sensitivity to verb biases.

(4) Tickle the pig with the fan (instrument biased verb)

(5) Choose the cow with the stick (modifier biased verb)

Interestingly, whilst biases in adults were found to interact with the referential context (they are still more likely to select a modifier interpretation in a two-referent-context than when there is only one referent), children's responses were entirely affected by verb bias irrespective of context suggesting that verb biases are a far more salient cue for children than the referential scene.

Additional evidence of the potency of verb biases (and of children's ability to demonstrate referential sensitivity in the absence of these biases) has been provided by Hurewitz, Brown-Schmidt, Thorpe, Gleitman and Trueswell (2000). Five year old children took part in a production task designed to elicit restrictive modifiers. A short

story (acted out by the experimenter) referred to pairs of animals (e.g., two frogs) taking part in different events followed by a specific question which required a restrictive modifier (e.g., *Which frog visited Mrs Squid's house?*). Under these circumstances children produced appropriate restrictive modifiers (e.g., *the frog on the napkin*) around 70% of the time. The same children were then required to immediately perform the *put* task used in Trueswell et al.'s (1999) original experiment. These children, who had just demonstrated their ability to take account of referential information in their utterances, performed incorrect actions at a similar level to those in the original Trueswell et al. (1999) study. As these children demonstrably possess the necessary knowledge of both the syntax and the situational constraints (and thus the referential principle) the most compelling explanation for these paradoxical data is that the verb bias associated with *put* is such a potent constraint for the young parser that it over-rides the information provided by the visual and discourse context.

In contrast Meroni and Crain (2003) have suggested that young children can demonstrate referential sensitivity with the constructions used by Trueswell and colleagues, and that 5-year-olds responses can approach adult levels when two straightforward situational modifications are made to the act-out task. They claimed that children's performance on the *put* task may be partly explained as an *order of mention* effect. Children may begin to plan (and possibly even execute) a response before they have fully compiled all the necessary information, and this can result in children acting out sentences in the order of mention rather than in the correct conceptual order. In the case of prepositional phrase sentences the suggestion is that children may begin planning to move the frog to the empty napkin before hearing the entire sentence and are then unable to inhibit this plan. Order of mention effects have

often been found in young children (Armidon & Carey, 1972; Clark, 1971; Hamburger & Crain, 1982) but they can be eliminated under conditions where the visual array is withheld until after the presentation of the target instruction (Matthei, 1982). This may prevent children from compiling an erroneous representation which they cannot later over-ride in the light of conflicting information.

Furthermore, based upon their observation that the incorrect frog is selected 90% of the time in the Trueswell et al. (1999) two-referent context Meroni and Crain (2003) suggest that the children may have made a pragmatic inference that the frog that was not on a napkin was the intended referent. Having done this they then assigned a destination role to the prepositional phrase *on the napkin*. This proposition explains children's non adult-like responses without suggesting that they lack sensitivity to the referential principle, because under this suggestion one frog is more salient therefore children have not in fact violated the referential principle (Meroni & Crain, 2003). The possibility of such pragmatic inferences has also previously been raised by Hurewitz et al. (2000) who controlled for them by placing both frogs on platforms (e.g., one on a chair and one on a toy tree) and found this did not change the overall parsing preferences of children. In contrast, Meroni and Crain (2003) found that 3- to 5-year-olds exhibited adult-like performance (around 90% correct) in a two-referent context when both frogs were placed on different colored napkins (referred to as a pragmatic block). Delayed responding was also enforced by asking children to close their eyes whilst listening to the sentence. Meroni & Crain (2003) claimed that these changes enabled children to inhibit their incorrect syntactic plans.

In summary, whilst previous findings strongly implicate the important role of verb bias and frequency based information for the young parser, Meroni and Crain's work implies that in certain circumstances such biases can be modified and over-

ridden by referential information. The relative importance of different information sources is crucial in terms of understanding the development of the cognitive hierarchy underpinning ambiguity resolution and ultimately comprehension.

Meroni and Crain's (2003) results warrant further exploration and replication as no direct comparison was made between the eyes-closed condition and an eyes-open condition, or between one-referent and two-referent-contexts (only two-referents were used, and results were compared to the original Trueswell at al., 1999 study). Moreover, as the pragmatic block and eyes-closed manipulations were introduced simultaneously, it is not clear whether either one or both of these manipulations together account for their findings. If the pragmatic block accounts for the effect it would suggest that contextual cues are in fact more salient to young children than adults, and that any violation of pragmatic expectations dramatically reduces their ability to resolve syntactic ambiguities. However, if the eyes-closed manipulation accounts for the improvement we may conclude that children need to be enabled to inhibit incorrect possibilities in order to access the correct representations.

The study reported below assesses children's comprehension of prepositional phrase ambiguities in both one- and two- referent scenes by directly comparing the experimental settings used by Trueswell, et al. (1999) with those adopted by Meroni and Crain (2003). The pragmatic-block and eyes-closed manipulations are also introduced separately. Eight- and eleven-year-old children are tested in addition to five-year-olds in order to track the development of syntactic processing with reference to contextual factors.

Method

Participants

Sixty children were recruited from participating primary schools in the Nottinghamshire area, UK, and were divided into three age groups (20 children per group); these were 5-year-olds (range 4;7 – 5;6; mean 5;1); 8-year-olds (range 7;7 – 8;6; mean 8;1) and 11-year-olds (range 10;7 – 11;6; mean 11;1). The two youngest age groups comprised equal numbers of male and females. The gender split was not quite equal in the oldest age group due to sampling issues (7 females; 13 males). All participants were monolingual English speakers raised in English speaking households and possessed no known language or hearing difficulties.

Materials

Test sentences were either based upon those used by Trueswell et al. (1999; e.g., sentence 6) or by Meroni and Crain (2003; e.g., sentence 7), and all were ambiguous prepositional phrase sentences as shown in the examples below. All sentences began '*Put the…*' followed by the name of the target animal, a prepositional phrase describing the location of the target, and a second prepositional phrase describing the intended destination of the target. As outlined above, the first prepositional phrase is temporarily ambiguous as to whether it is intended as a modifier (describing a particular 'frog') or as a destination (for the 'frog' to be moved to).

(6) Put the frog on the napkin in the box

(7) Put the frog on the red napkin in the box

The prepositions used in the first prepositional phrase were *on* and *in* and *with*, and for the second prepositional phrase *on*, *in*, *behind* and *under*. Unlike Trueswell et al. (1999), unambiguous controls were not used, as in this study the comparison of interest was between modes of presentation, rather than ambiguous/unambiguous sentence type[1]. A full list of the sentences used is provided in Appendix A.

Modes of presentation

The experiment comprised four different presentation conditions with 12 sentences being presented in each of the following scenarios:

TSHL Condition

This condition was intended as a replication of Trueswell, Sekerina, Hill & Logrip's (1999; here abbreviated to TSHL) original effect with 5-year-olds, further investigating the developmental trajectory of this effect. In this condition sentences closely followed those used by Trueswell at al. (1999) and only the target animal was placed upon a platform (e.g., a napkin). For examples of the visual arrays used see pictures 1 and 2 in Figure 1.

TSHL-eyes-closed Condition

This condition used Trueswell et al. (1999) style sentences and arrays, but adopted the procedure advocated by Meroni and Crain (2003), whereby children were asked to close their eyes and face away whilst listening to the sentences.

Pragmatic Block Condition

This condition used sentences like those used by Meroni and Crain (2003). In this case both the target and distracter animal were placed on visually distinct platforms (e.g., different colored napkins, as shown in picture 3, Figure 1) and this platform was referred to in the initial preposition (e.g., *put the frog on the red napkin*…). In all other respects this condition followed Trueswell et al. (1999).

M&C Condition

This condition was a complete replication of Meroni and Crain's (2003; here abbreviated to M&C) procedure. In this condition pragmatic block sentences were used, along with the eyes-closed procedure.

Within each of these presentation conditions half the trials were one-referent trials, including only one target animal (e.g., one frog and one duck), and the other half were two-referent trials including two identical target animals (e.g., two frogs) with the target being identified by its platform (e.g., napkin). This yielded 48 trials per child (i.e., 12 per presentation scenario with 6 trials in each scenario being either one- or two- referent).

Presentation type and number of referents were fully crossed with all participants experiencing both one- and two-referents in each of the four presentation scenarios. The four presentation scenarios were blocked and fully counterbalanced in an attempt to control for any strategic carry-over effects from one condition to another. One- and two-referent scenes were equally distributed within these blocks.

Procedure

An act-out procedure was used so children were asked to move toy animals in accordance with the experimenter's instructions. The exact instructions given can be found in Appendix B. Before commencing the experiment children were asked to verbally identify all of the animals and props in order to establish that they recognized the objects to be used. In cases where a different name was used to label any of the props that name was adopted by the experimenter for the duration of the experiment. As a practice the children were asked to follow the experimenter's instructions to move some of the objects. All participants were able to recognize the objects and understood the task during the practice session.

For each trial participants were positioned in front of a table upon which the array of toys was placed. In trials under the *TSHL* and *pragmatic block* conditions children were facing the table whilst listening to the sentence, and accordingly the instruction began with "now look at the table". However, in the two *eyes-closed* conditions children were positioned facing away from the table when hearing the sentence and consequently the instruction began with "now look at me" (referring to the experimenter). They were then instructed to turn round to perform the action after hearing the sentence. Each child acted out all 48 sentences in succession with short breaks between each trial whilst the experimenter changed the visual display. When each scene was changed the props used in the previous scene were completely removed from view to avoid any confusion for the child about what they should be attending to. The same female experimenter read all the sentences with conscious effort being made to retain continuous, neutral and unstressed prosody. Children were given positive feedback after each trial and all participants were given a sticker on completion of the experiment as a reward for taking part.

Results

Scoring and Coding of Results

For each trial responses were carefully coded by the experimenter closely following Trueswell et al.'s (1999) coding scheme. Act-out responses were scored correct if participants moved the target animal to the correct destination without moving any animals to the incorrect destination. To illustrate for sentence (7) above the correct response would be to move the frog that is already on a napkin into the box[2]. Incorrect responses were then classified into different error types. Further details of the errors that occurred are discussed below. First, the pattern of correct responses will be considered.

Analysis of correct act out responses

The proportion of correct responses in each condition for each child was calculated. The mean percentages of correct trials are given in Table 1. All statistical analyses were carried out on the arcsine-transformed data (the untransformed data revealed the same patterns).

The pattern of results shown in Table 1 suggests that the youngest children were often unable to avoid the erroneous destination interpretation when there were two referents. Indeed the data suggest that 5-year-olds made misinterpretations more often in the two-referent condition than when there was just one referent. Older children correctly acted out the sentences most of the time with a less pronounced effect of referents. The data shown here for 5-year-olds are broadly consistent with the earlier findings of Trueswell et al. (1999). There was an increase in correct actions as a function of age but little effect of type of experimental scenario with performance in two-referent-contexts being consistently lower than in one-referent-

contexts across all conditions for 5-year-olds, and lower or roughly equal across all conditions for 8- and 11-year-olds.

In order to investigate these results statistically, subject means were entered into an ANOVA with two within-subjects factors: Mode of presentation (TSHL, TSHL-eyes-closed, Pragmatic block, M&C), and Number of Referents (one-referent, two-referent), and the between subjects factor of age group (5-, 8- and 11-years-old). This revealed a highly significant effect of number of referents, $F(1, 57) = 21.401$, $p < .0001$, partial $\eta^2 = .273$, with correct actions being performed more often when there was just one referent. There was also a highly significant main effect of age, $F(2, 57) = 18.139$, $p < .0001$, partial $\eta^2 = .247$, and performance improved with age. There was no main effect of mode of presentation ($F < 1$).

In order to further investigate the main effect of age a TUKEYS HSD was performed. This revealed that 5-year-olds differed significantly from 8-year-olds ($p = .002$), and 11-year-olds ($p = .001$), but that 8- and 11-year-olds did not significantly differ (both $p > .9$). It is evident that the main difference between the young children (5-year-olds) and older children (8- and 11-year-olds) arose from the generally lower proportion of correct actions made by young children, especially when there were two potential referents. Evidently for this age group correct responses were reduced as a function of having to select between two identical referents. As in Trueswell et al. (1999) these data suggest that the modifying prepositional phrase (*on the napkin*) "was rarely taken as a Modifier, resulting in chance performance when determining a referent for the direct object NP '*the frog'*." (Trueswell et al., 1999, p.106). There was no evidence in this age group that manipulating the experimental scenario improved performance, and certainly no evidence that it resulted in the correct interpretation of *on the napkin* as a modifier. Children of this age appear not to utilize

the information in the referential scene to disambiguate the sentence, irrespective of mode of presentation.

Inspection of Table 1 reveals that by around the age of 8 children performed significantly better than 5-year-olds as correct actions occurred above 70% of the time in all conditions (compared with between 31% and 56% for 5-year-olds). By this age the children regularly interpreted *on the napkin* as a modifier; the visual scene did not seem to influence this interaction although even by the age of eleven children still found it slightly easier to make the correct actions when there was only one potential referent to contend with.

Mode of presentation was not found to interact with age ($F < 1$) but an interaction between number of referents and age was found to approach significance, $F(2, 57) = 2.971$, $p = .059$, partial $\eta^2 = .094$. Inspection of the means indicates that this interaction arose because 5-year-olds made many more correct responses when there was one-referent compared with two-referent trials, whereas this difference was less pronounced for the older children.

There was a significant presentation by referent interaction, $F(3, 57) = 3.015$, $p = .032$, partial $\eta^2 = .050$, but no three way presentation x referent x age interaction, $F(6, 57) = 1.682$, $p > .1$, partial $\eta^2 = .056$). In order to investigate the interaction between presentation and number of referents post hoc paired t-tests were conducted between one- and two- referent conditions within each mode of presentation. It was found that whilst the number of correct actions was significantly higher in a one-referent condition for both the TSHL, $t(59) = 3.709$, $p < .0001$, and TSHL-eyes-closed, $t(59) = 3.980$, $p < .0001$, conditions this was not the case for the pragmatic block, $t(59) = 1.851$, $p = .069$, or the M&C, $t(59) = 1.328$, $p > .1$, conditions. However it is interesting to note the correct actions in the two-referent context with

the pragmatic block and M&C scenarios did not differ significantly from the two referent TSHL condition, $t < 1$; $t(59) = 1.069$, $p > .1$, respectively. This undermines the suggestion that the pragmatic block simply increased salience in a two referent condition as in this case performance should improve under conditions employing this manipulation. Overall, the data suggest that children performed correct actions at similar levels irrespective of mode of presentation[3].

Comparison of early and late trials

These results did not replicate Meroni and Crain's (2003) findings, even under the condition designed to mirror their study. However, we presented children with multiple conditions; they had to switch between one- and two- referent contexts[4], as well as between the different modes of presentation. In contrast participants in Meroni and Crain's (2003) experiment were only exposed to two- referent contexts under one mode of presentation. Exposing children to multiple conditions may have de-emphasized the informativeness of the visual scene in the present study. That is, children may have become aware that the visual scene was *not always* useful in disambiguating sentences and decided to ignore it. In this case we may observe evidence of sensitivity to context during early trials, which diminishes later in the experiment. In order to investigate this, trials from the first block for each participant were compared with trials from the last block. The means presented in Table 2 revealed that 5-year-olds performed poorly even during the first block and an overall decrease in performance was observed between blocks. This decrement was particularly marked when there were two referents (46% correct actions in the first block compared with 25% in the last block). Older children seemed to perform in a relatively stable way across the experiment.

An ANOVA with the two within-subjects factors of block (first or last) and number of referents (one-referent, two-referent) and the two between-subjects factors of age group (5-, 8- and 11-years-old) and experimental group was conducted. This revealed that the decrement was significant given the main effect of block, $F(1, 48) = 6.282$, $p = .01$, partial $\eta^2 = .116$. Consistent with the main analysis a significant main effect of referents was also observed, $F(1, 48) = 11.361$, $p = .001$, partial $\eta^2 = .191$, reflecting the advantage for one-referent trials. However there was no block by referent interaction ($F < 1$) suggesting that performance decreased in later trials irrespective of number of referents. A main effect of age group was also observed, $F(2, 48) = 6.667$, $p = .01$, partial $\eta^2 = .265$, which was found to interact with block, $F(2, 48) = 3.326$, $p = .044$, partial $\eta^2 = .122$, but not referents, $F(2, 48) = 1.768$, $p = .182$, partial $\eta^2 = .069$.

The age x block interaction may be explained by differences between the 5-year-olds and the two older age groups because the means suggest that whilst the young children performed less well in the last block compared to first the older children's performance was consistent across blocks. A decrement was particularly evident for 5-year-olds in the two-referent condition. A post hoc ANOVA confirmed that there was a main effect of block for 5 year olds alone, $F(1, 16) = 7.512$, $p = .015$, partial $\eta^2 = .320$. The data from the 8- and 11-year olds was combined into one further post hoc ANOVA (given the similarity in the means for these groups)[5] and confirmed that there was no main effect of block for the older children, $F(1, 36) < 1$. A main effect of number of referents was found for the 5-year-olds, $F(1, 16) = 10.531$, $p = .005$, partial $\eta^2 = .397$, and approached significance for the older children, $F(1, 36) = 3.389$, $p = .074$, partial $\eta^2 = .086$, suggesting that younger children demonstrated a more reliable advantage for the one-referent scenes. Reflecting the

main analysis a block x referent interaction was not statistically significant for 5-year-olds ($F < 1$) indicating that their performance decreased in the last block in both 1- and two-referent conditions. However, the striking differences between means obtained with this age group were subjected to planned t-tests which revealed that the difference between blocks was significant for two-referent scenes, $t(19) = 3.253$, $p = .04$, but not for one-referent scenes, $t(19) = 1.337$, $p = .197$. This indicated that 5-year-olds' performance decreased more markedly when there were two referents.

This analysis provides no evidence that that 5-year-olds successfully used referential context to resolve the ambiguity even in the first block of the experiment, they choose either frog equally often in a two-referent-context. By the last block children chose the incorrect frog 75% time suggesting that they were biased in some way to select the incorrect frog. It seems that these young children may have begun by selecting either frog at chance but then adopted a strategy of interpreting the next viable information as the anticipated destination. Therefore in the Trueswell conditions (only one frog on a napkin, other frog not on a platform) they identified the empty napkin as the destination (upon hearing "napkin") and decided early on that the frog not already on a napkin was to be moved. Children in the other conditions (both frogs on napkins) similarly assumed that the empty napkin of the color mentioned (e.g., the red one) was the destination and that the frog on the other colored napkin should be moved. In each case the end result is the same – the incorrect animal will be selected and moved unless this initial plan is revised.

Experimental group was also entered into the ANOVA to investigate whether exposure to any one mode of presentation during the first block differentially affected performance on the last block. However, no main effect of experimental group was found ($F < 1$) and all other interactions were non significant.

Error data

Incorrect responses were classified into different error types. Descriptive error data are provided here to elucidate the main statistical analyses given above. Most of the observed error types had previously been noted by Trueswell et al. (1999) and were synonymously labeled as the *hopping* error, the *falling short* error and the *one of each error*.

The most common error across children of all ages was the hopping error involving the movement of one of the animals first to the incorrect destination and then to the correct destination (this also sometimes included moving the empty napkin to the correct destination). Another common error was the falling short error noted when the child ignored the content of the second prepositional phrase (e.g., *in the box*) placing either the target animal or the other animal on the incorrect destination and moving nothing to the correct destination. The one of each error involved the performance of two separate actions; first the child moved one animal to the incorrect destination and then the other animal to the correct destination. Whilst this latter error was observed in our data it was rare.

Other errors noted in the current study were the *Incorrect animal* error, which involved simply moving the incorrect animal to the correct destination, the *Up & Down* error which was the simple action of lifting the target animal up but then putting it back on its original destination; and the *Both Animals* error involving the action of putting both animals into the correct destination. Inspection of these errors revealed that there was no systematic variation in error types according to mode of presentation. Production of these errors did however vary with age, in that both 8- and 11-year-olds primarily made hopping errors (accounting for 16% of all responses

by each of these groups) and rarely made any of the other error types (the total of all other errors accounted for only 3% and 2% of total responses for the 8- and 11-year olds, respectively). Hopping was also the most common error observed in the 5-year-olds (accounting for 41% of total responses made by this age group), followed by the falling short error (accounting for 5% of responses), and with several children also performing the other errors in the other categories. Figures 2 and 3 illustrate the types of errors made by 5-year-olds as a function of mode of presentation, in one-referent and two-referent conditions, respectively. The older children's errors have not been illustrated due to the relatively small number of errors observed in this age group and the homogeneity of error types (as noted). Importantly errors do not appear to be mediated by changes to the mode of presentation and 5-year-olds continue to demonstrate a robust inability to utilize two-referent contexts across experimental manipulations.

Like Trueswell et al. (1999) we observed that the object which was selected first (in the two referent conditions) was closely linked with whether the correct destination was selected or not. It was found that the correct animal was moved to the correct destination on 68% of trials. In contrast the incorrect animal was moved directly to the correct destination on only 3% of trials. Initially selecting the correct *frog* seems to be a precursor to performing the subsequent correct actions. This point will be discussed further in section 4.

4. General Discussion

To summarize, these results replicate Trueswell at al.'s (1999) finding that five year old children struggle to interpret ambiguous prepositional phrases as a modifier, even in a two-referent-context which supports the need for modifying information.

This also holds true when mode of presentation is varied to discourage incorrect pragmatic inferences and ballistic responding. The results suggest that an adult-like pattern emerges by around the age of eight years, with most children at this age and above demonstrating the ability to resolve these ambiguous sentences successfully. Older children's performance in this study can be taken as evidence that they make use of the referential principle to resolve syntactic ambiguity. Although they do not do *better* in a two-referent-context compared with a one-referent-context (as might be predicted by the referential theory) performance in this age group reaches at least 70% correct in all conditions therefore approaching ceiling. Trueswell at al. (1999) found adults did not perform more correct actions in a two-referent-context either (they were also at ceiling with performance above 90%) but examination of adults' eye movements revealed sensitivity to referential context. In a one-referent-context eye movements revealed that the empty napkin was considered as a destination for the frog during the ambiguous prepositional phrase, whereas in a two-referent-context this effect was eliminated with adults rarely looking to (and therefore considering) the empty napkin. On this basis we would expect to see a similar pattern of eye movements emerging in children around the age of eight years old (Trueswell et al., 1999 report preliminary findings to this effect). Crucially Meroni and Crain's (2003) findings have not been replicated. Five-year-olds did not utilize information from the visual scene to resolve syntactic ambiguity and children's inability to demonstrate referential sensitivity with prepositional phrase ambiguities is affirmed.

It seems that children decide which object to move at a very early stage of processing, and that they are unable to revise this commitment in the light of the incoming language. This occurs even in situations where the planning of a response is delayed until after the sentence has been heard in its entirety. Trueswell at al.

(1999) observed that when the incorrect animal was selected it was usually moved to an incorrect destination, but that when the correct animal was moved it was usually moved to the correct destination. It was suggested that this may indicate that children make an incorrect pragmatic inference that the frog not already on a napkin is the one to be moved (Trueswell et al., 1999, Meroni & Crain, 2003 but see Hurewitz et al., 2000). We observed the same pattern even in conditions designed to prevent this inference (both frogs were already on napkins). This finding is consistent with Hurewitz at al. (2000) who demonstrated that the introduction of additional platforms did not eliminate the ambiguity effect with 5-year-olds. As suggested by Trueswell, Papafragou & Choi (in press; see also Novick et al., 2005) this inability to revise may be related to the development of inhibitory control, and possibly to limited working memory span (in support of this there is evidence that adults with limited working memory capacity are also unable to revise incorrect syntactic commitments, Mendehlsohn, 2003).

Given the size of the sample and number of items used in the present study, and also that the results so clearly mirror those of Trueswell et al. (1999) and others (Hurewitz et al., 2000), the observed effect is robust and replicates across a variety of experimental manipulations. Therefore this study supports the conclusions put forward by Trueswell in his original paper (Trueswell et al., 1999), and since (Snedeker & Trueswell, 2004; Trueswell et al., 2006). Evidence from studies using the act-out task strongly suggests that whilst children demonstrate sensitivity to referential context under some circumstances (e.g., Hurewitz et al., 2000) it is easily overridden by other constraints, including strong verb biases when presented with verbs like "put" (Snedeker & Trueswell, 2004).

This conclusion supports constraint-based models of parsing, where context is viewed as one possible factor competing with other information sources in order to arrive at a syntactic interpretation. The errors made by young children can be seen as arising from the experiential development of the parser. Initially constraints based upon the probabilistic facts of the language, such as verb biases are the most reliable constraint available to the child, and hence are favored by the developing parser. As further knowledge of the situational contexts within which certain utterances may occur is encountered this information source becomes statistically more reliable to the child and develops to become a stronger constraint.

The results of the present study clearly suggest that the improvement in performance reported by Meroni and Crain (2003) cannot be attributed to their experimental manipulations (pragmatic blocking and delayed responding), so we can only suppose that other aspects of their design improved performance. They report in their procedure that "children were told a short story about the events that led up to a particular arrangement of characters and props" (Meroni & Crain, 2003, p12). Evidence from the wider act-out literature has suggested that placing the act out task in a discourse context can drastically improve act out performance (Correa, 1995; Kidd & Bavin, 2002; Weighall & Altmann, 2001; Weighall, 2003). However, this is unlikely to fully explain their observed findings. Recall that Hurewitz et al. (2000) placed "put" sentences in just such a discourse context and found that, although children clearly demonstrated sensitivity to the referential principle in their production, performance on the "put" act-out task did not improve. Based upon these findings it is suggested that adult-like performance in 5-year-olds would not emerge that even if the current study was replicated with a discourse condition (indeed pilot data recently collected in our lab confirms this). However, it is possible that all of

these manipulations may have an additive effect such that combining pragmatic block with delayed responding in a discourse context may provide children with the best possible chance of succeeding with this particular task.

The effect of switching between number of referents and modes of presentation within one experiment was also explored as a possible explanation of the difference between studies. It seemed plausible that children may have learned that context was not a reliable cue to disambiguation in this particular task, thus deciding to ignore it. However, there was no evidence that 5-year-olds were able to use referential information in the early two-referent trials of the experiment with performance being at chance (46%). Five-year-olds' performance decreased to below chance (25%) by the last block and this seemed to reflect the fact that they routinely interpreted the first prepositional phrase as the intended destination and then adopted a strategy of moving the animal not already on that destination (e.g., upon hearing red napkin the frog not already on a red napkin would be moved to the empty red napkin). Whatever else is going on here it seems certain that young children readily make a destination interpretation of the first prepositional phrase and do not make use of referential context to avoid or revise this initial interpretation. Not only do the children fail to revise this interpretation on each individual trial they fail to revise it over repeated trials. Having misinterpreted the first prepositional phrase they often interpret the second prepositional phrase (in the box) as a subsequent destination hence the high proportion of hopping errors.

One explanation for young children's reluctance to revise their initial plan even upon hearing "in the box" may be that they adopt a bird in the hand strategy (Legum, 1975). This strategy arises because children are often reluctant to put down a toy they have selected once they have picked it up. Hence they adopt the strategy of

making the toy-in-the–hand do all the incoming actions – this phenomenon has been found to be prevalent in act out tasks investigating other complex, but non ambiguous, structures (such as sentences containing a relative clause e.g., Hamburger & Crain, 1982; 1984).

Meroni and Crain (2003) suggest that a bird-in-the-hand strategy may explain the observation that initially selecting the correct animal seems to be a precursor to arriving upon the correct syntactic interpretation (evident in Trueswell et al., 1999, and in the data presented here). If we accept that once children have selected a "frog" (as described above) they will then persevere with making the frog do the required actions, then upon selecting the correct frog (already on a napkin) the only sensible actions available based upon the input are to move that frog to the empty napkin (a common error in 5-year-olds), or to move it to the empty box (thus landing upon the correct response). Whereas if the incorrect frog (not already on a napkin) is selected the next available destination to be mentioned in the linguistic input is the empty napkin. This was also a common error. Furthermore evidence for the bird-in-the-hand strategy can also be seen in the eye movement record reported in Trueswell at al. (1999); the object that is fixated upon first (the 'bird-in-the-hand') is the one most likely to be picked up first, and is then the one that is used to act-out the sentence. Importantly, this is only true for ambiguous structures; the same dependency was not found by Trueswell et al. for the unambiguous alternative "*Put the frog <u>that's</u> on the napkin...*". This suggests that by age five children understand the function of the complementizer *that*, thus correctly interpreting *on the napkin* as a modifier. When there is no ambiguity children construct an accurate representation of the sentence, resulting in actions that are independent of which object was initially fixated - if they happened to fixate the incorrect animal they shift this fixation to the correct one upon

hearing the modifier. However, when the ambiguous sentence is encountered a strong destination bias is revealed, coupled with increasingly strategic responses based upon which target had the child's attention first.

To summarize, the evidence presented here clearly and robustly replicates Trueswell at al.'s (1999) finding that 5-year-olds prefer a destination interpretation when faced with this type of ambiguity, and do not utilize referential information to over-ride this preference. The constraint-based learner account provides a valuable explanation of the mechanisms underlying such non-adult performance. It is concluded that verb biases play a potent role in young children's ambiguity resolution, and that referential context begins to interact with (and over-ride) such biases as children's knowledge of situational constraints and discourse context increases such that adult-like processing emerges at around the age of eight years old. However, given the evidence of strategic responding in the data presented here, an important question arises as to how these findings will translate across different tasks. Whilst it seems likely that children's verb biases would be evident irrespective of task perhaps children might demonstrate a greater ability to revise their interpretation if presented with a task that was less open to such a strategic effect. This issue will require further empirical investigation.

Acknowledgements

I thank Michelle Thompson for collecting the data presented here as part of her final year undergraduate research project. I thank two anonymous reviewers for their helpful comments; and Jane Morgan, and Gerry Altmann for their support and advice; and John Reidy for advice about the interpretation of effect sizes.

References

Altmann, G.T.M (1998). Ambiguity in Sentence Processing. *Trends in Cognitive Sciences, 3,* 146-152

Altmann, G.T.M., Garnham, A. & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language, 31,* 685-712.

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

Armidion, A. & Carey, P. (1972). Why five year olds cannot understand before and after. *Journal of Verbal Learning and Verbal Behaviour, 11,* 412-423.

Bates, E., Devescovi, A. & D'Amico, S. (1999). Processing complex sentences: A cross-linguistic study. *Language and Cognitive Processes, 14,* 69-123.

Britt, M.A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. *Journal of Memory and Language, 33,* 251-283.

Clark, E.V. (1971). On the acquisition of the meaning of *before* and *after*. *Journal of Verbal Learning and Verbal Behavior, 10,* 266-275.

Correa, L.M. (1995). An alternative assessment of children's comprehension of relative clauses. *Journal of Psycholinguistic Research, 24,* 183-203.

Eberhard, K., Spivey-Knowlton, M., Sedivy, J. & Tanenhaus, M.K. (1995). Eye movements as a window into real-time spoken language processing in natural contexts. *Journal of Psycholinguistic Research, 24,* 409-436.

Ferreira, F. & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348-368.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading*, Hillsdale, NJ: Erlbaum.

Frazier, L. (1989). Against lexical generation of syntax. In W.D. Marslen-Wilson (Ed.), *Lexical Representation and Process*. Cambridge, MA: MIT Press.

Frazier, L. & Fodor, J.D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291-325

Goodluck, H. (1990). Knowledge integration in processing and acquisition: Comments on Grimshaw and Rosen. In L. Frazier & J. De Villiers (Eds.), *Language Processing and Language Acquisition.* Dordrecht: Kluwer.

Goodluck, H., & Tavakolian, S. (1982). Competence and processing in children's grammar of relative clauses. *Cognition, 11,* 1-27.

Hamburger, H. & Crain, S. (1982). Relative acquisition. In S. Kuczaj (Eds.), *Language Development II.* Hillsdale, NJ.: Lawrence Erlbaum.

Hamburger, H. & Crain, S. (1984). Acquisition of cognitive compiling. *Cognition, 17,* 85-136.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Trueswell, J.C. & Gleitman, L. (2000) One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research*, 29, 597-626

Kidd, E. & Bavin, E.L. (2002). English-speaking children's comprehension of relative clauses: Evidence for general cognitive and language specific constraints on development. *Journal of Psycholinguistic Research, 31,* 599-617.

Legum, S. (1975). *Strategies in the acquisition of relative clauses*. Southwest Regional Laboratory Technical Note, No. TN 2-75-10.

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

Mendelsohn, A. (2003) The role of inhibition in sentence processing as demonstrated by a new task. Paper presented at the *CUNY Conference on Human Sentence Processing*, Cambridge, MA, March 2003.

Meroni,L. & Crain, S. (2003). On not being led down the kindergarten path. *Proceedings of the 27th Boston University Conference on Language Development*, 531-544, Cascadilla Press, Somerville, MA.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience, 5*, 263-281.

Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behaviour*, **13**, 272-281.

Snedeker, J. & Trueswell, J.C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, 49, 238-299.

Spivey, M.J., Tanenhaus, M.K., Eberhard, K.M. & Sedivy, J.C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, 45, 447-481

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M. & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632-1634.

Tavakolian, S. L. (1981). The conjoined-clause analysis of relative clauses. In S. L. Tavakolian (Eds.), *Language acquisition and linguistic theory*. (pp. 167-187) Cambridge, Mass.: MIT Press.

Trueswell, J.C. & Gleitman, L. (2004) Children's eye movements during listening: Developmental evidence for a constraint-based theory of sentence processing. In J.M. Henderson & F. Ferreira (Eds.), *The Interface of Language, Vision, and Action: Eye Movements and the Visual World.* New York: Psychology Press.

Trueswell, J.C., Papafragou, A. & Choi, Y. (in press). Referential and Syntactic Processes: What Develops? To appear in E. Gibson and N. Pearlmutter (Eds.) *The Processing and Acquisition of Reference.* Cambridge, MA: MIT Press.

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

Weighall, A. (2003). Integration between language and visual context: A re-investigation of children's interpretation of relative clauses. *Unpublished PhD Thesis, University of York, U.K.*

Weighall, A. & Altmann, G.T.M. (2001). When two cats are better than none: Children's interpretation of relative clauses (revisited). Paper presented at the *7th Annual Conference on Architectures and Mechanisms for Language Processing (AmLAP),* Saarbrucken, Germany, September, 2001.

Appendix A

1. Put the fish on the sponge on the plate
2. Put the bird on the road in the cage
3. Put the hamster on the napkin in the box
4. Put the cat in the cup on the saucer
5. Put the horse in the field in the stable
6. Put the rabbit on the towel in the cage
7. Put the dog with the stick on the tray
8. Put the snake in the tunnel in the bowl
9. Put the cow on the saucer behind the fence
10. Put the lamb on the soap behind in the field
11. Put the sheep on the plate behind the wall
12. Put the pig in the tunnel with the tree
13. Put the horse on the towel in the box
14. Put the pig in the pond on the jug
15. Put the donkey on the plate in the wagon
16. Put the guinea pig on the road in the teapot
17. Put the rabbit in the bowl in the tunnel
18. Put the snake behind the fence on the sponge
19. Put the dog in the woods with the stick
20. Put the bird in the cup on the towel
21. Put the fish in the pond on the napkin
22. Put the sheep in the jug in the tunnel
23. Put the cow in the field on the floor
24. Put the cat on the napkin in the basket

25. Put the dog with the red napkin behind the wall
26. Put the horse on the white plate in the stable
27. Put the fish with the blue towel in the pond
28. Put the sheep in the large field in the woods
29. Put the lamb on the stripy circle behind the fence
30. Put the bird in the black tunnel in the cage
31. Put the pig on the spotty square in the tent
32. Put the cat in the white cup on the towel
33. Put the guinea pig on the green triangle on the square
34. Put the cow in the large box in the truck
35. Put the horse on the short road in the wagon
36. Put the snake in the small tunnel on the tray
37. Put the guinea pig with the stripy towel under the wagon
38. Put the bird in the yellow mug on the napkin
39. Put the sheep on the white saucer with the flag
40. Put the lamb with the pink napkin in the tent
41. Put the donkey on the spotty circle in the field
42. Put the rabbit on the small plate in the cage
43. Put the fish on the white soap in the box
44. Put the hamster with the square tub under the towel
45. Put the cat on the peach sponge in the truck
46. Put the dog on the long road in the teapot
47. Put the mouse in the dark tunnel on the saucer
48. Put the pig with the white flag in the canoe

Appendix B

Exact Instructions

'I am going to ask you to do some short tasks moving these toys and animals. I am going to put some objects out on the table and then I want you to listen very carefully to me while I tell you what I want you to do. Try your best to do what I ask you to do. When you are done I want you to tell me that you have finished, ok?'

Footnotes

[1] Fillers were not used either, as pilot testing suggested that the inclusion of fillers made the task too lengthy to maintain the crossed design – again this is a deviation from Trueswell et al.'s method, however given that our results closely mirrored theirs in several important ways we do not feel that these changes adversely affected our results.

[2] Following Trueswell et al. (1999) 'on modifier' responses where the correct frog *and* the napkin are both placed in the box were also classed as correct, as in such cases the first prepositional phrase has still be interpreted correctly as a modifier.

[3] Further analyses were also conducted to investigate whether children may differentially select the correct animal (irrespective of the subsequent actions) as a function of experimental scenario. These results revealed a very similar pattern to the main analysis and there was no effect of experimental scenario.

[4] This is also true of Trueswell et al. (1999) who similarly used a repeated measure design.

[5]Separate ANOVAs for each age group revealed the same statistical pattern.

Table 1

Mean correct act-outs (%) in each condition as a function of age. Standard deviations are in parentheses.

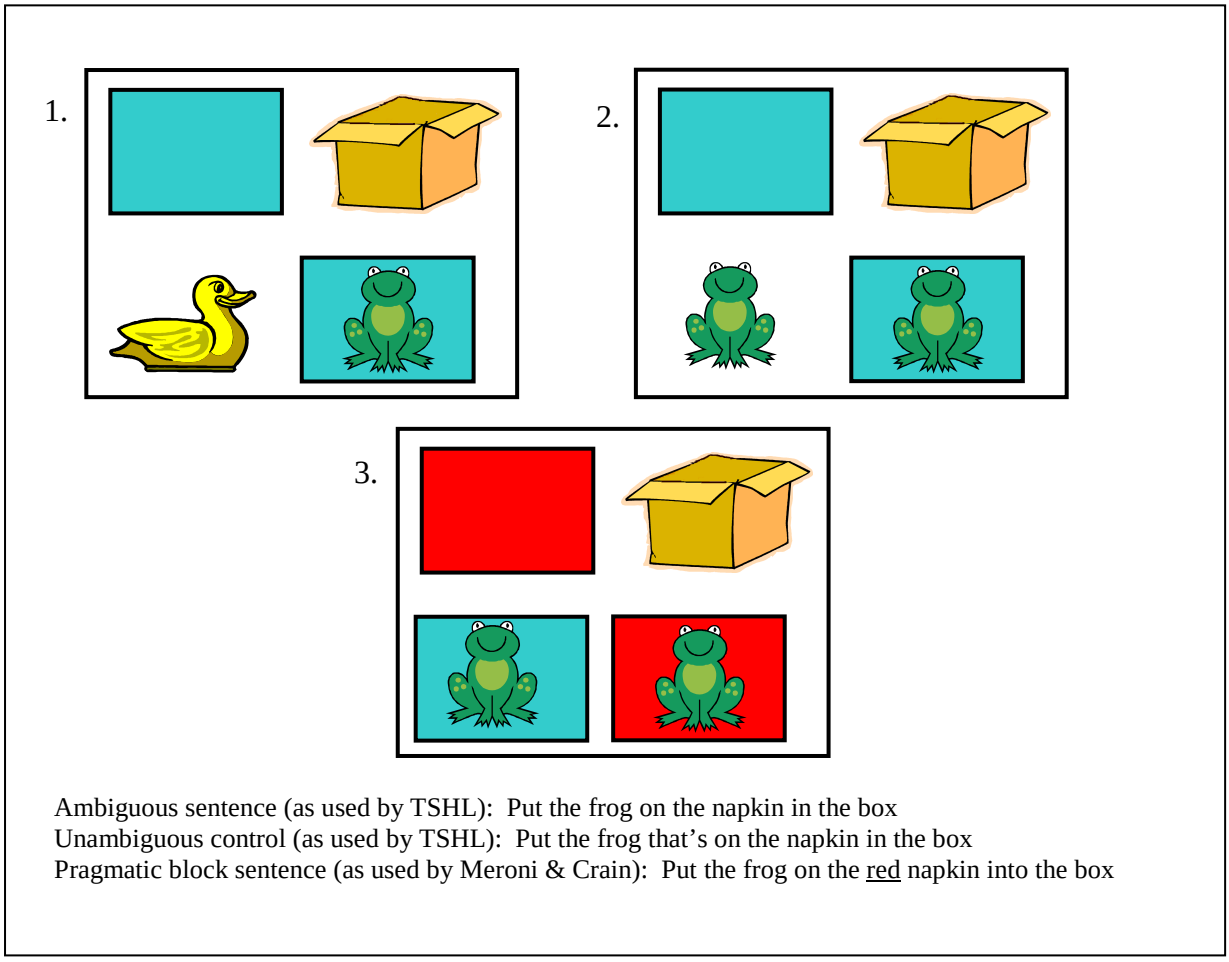
Scenario	TSHL		TSHL-eyes-closed		Pragmatic Block		M & C	
	1-ref	2-ref	1-ref	2-ref	1-ref	2-ref	1-ref	2-ref
5 years	50 (40)	34 (33)	56 (42)	31 (39)	44 (43)	38 (41)0	52 (40)	36 (44)
8 years	91 (18)	80 (32)	87 (26)	78 (36)	81 (33)	72 (41)	73 (36)	78 (39)
11 years	88 (31)	75 (38)	84 (36)	81 (37)	83 (31)	79 (33)	84 (36)	87 (31)
overall	76 (76)	63 (40)	76 (37)	63 (43)	69 (40)	63 (42)	70 (39)	67 (44)

Table 2

Mean correct act-outs (%) in the first and last block as a function of age and number of referents.

Scenario	First block		Last Block	
	1-ref	2-ref	1-ref	2-ref
5 years	55	46	45	25
8 years	82	75	79	78
11 years	88	81	83	78
overall	75	69	70	61

Figure 1
Examples of the array of objects used in a 1-referent (1) and a 2-referent (2) context in the Trueswell et al. (1999) study, and in the Meroni & Crain (2003) study (3).



Ambiguous sentence (as used by TSHL): Put the frog on the napkin in the box
Unambiguous control (as used by TSHL): Put the frog that's on the napkin in the box
Pragmatic block sentence (as used by Meroni & Crain): Put the frog on the red napkin into the box

Figure 2
Percentage of error types made by five year olds in each condition for **one** referent contexts

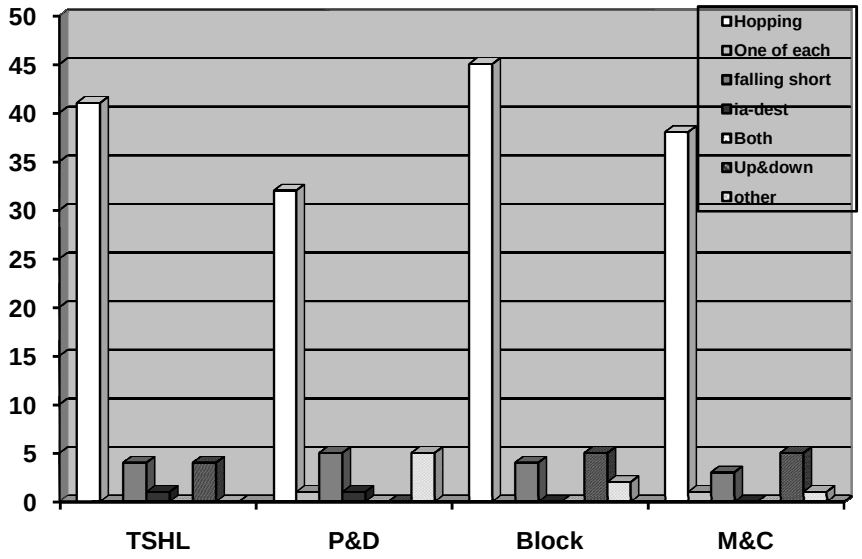


Figure 3
Percentage of error types made by five year olds in each condition for **two** referent contexts