Too Many Ways of Interpreting 'What Infants Know about Syntax'

Micha Y. Breakstone
Malka Rappaport Hovav
The Hebrew University of Jerusalem

1 Goal

In their influential paper Lidz, Waxman & Freedman (2003; henceforth LWF) argue that the acquisition of anaphoric 'one' in English can be explained only if innate syntactic knowledge is assumed. To support their claim LWF integrate both corpus and experimental evidence into the standardly structured argument for the poverty of the stimulus.

Following publication, LWF's work sparked a lively debate resulting in a host of commentaries which both challenged their conclusions and offered a variety of alternative accounts for the data, including those based on Bayesian learning models and pragmatically-driven learning. This paper critically reviews the exchange of critiques and the ensuing response and rejoinder spawned by LWF's work. We suggest that while LWF have successfully rebutted most of the critiques, an argument in one of the critiques remains valid. This critique suggests that learning is not necessarily contingent upon exposure to data which unequivocally resolves 'one' correctly, but rather that pragmatic considerations can facilitate acquisition of the relevant syntactic knowledge. We take this argument one step further and claim that the results reported by LWF can be explained without assuming syntactic knowledge altogether (other than the capability to distinguish between syntactic categories). We contend, furthermore, that LWF's argument that children possess knowledge of the hierarchical structure of NP, presupposes knowledge of hierarchy, and is thus circular. Our conclusion then is that LWF have *not* shown that infants possess the proposed syntactic knowledge.

We maintain, nonetheless, that LWF have successfully demonstrated that in the process of resolution of anaphoric 'one', infants do indeed rely on some *linguistic* knowledge. The nature of this knowledge – whether it is syntactic or of some other kind – remains to be determined. To this end we propose a follow-up experiment to distinguish among possible competing hypotheses. Also, along the way we present a refined characterization of the usage of anaphoric 'one'.

Crucially, it is not our goal to show that innate syntactic knowledge is not a prerequisite for language acquisition, but only that one must be extremely careful about evidence presented in support of such arguments, especially, when such arguments regularly come under attack by scientists employing general statistical models.

2 LWF's Argument for Innate Syntactic Knowledge

In their original paper, LWF present experimental results which they take as evidence that 18-month old infants possess knowledge of the internal phrase structure of NPs. Such knowledge, LWF argue, could not have been acquired by the infants from the linguistic environment, because the environment does not provide sufficient evidence, and should thus be taken as support for the existence of innate syntactic knowledge. LWF's argument, originally formulated in Baker, 1978 (see also Hamburger & Crain, 1984; Hornstein & Lightfoot, 1981; Lightfoot, 1982) is structured as follows.

It is generally accepted that NPs contain an internal *nested structure*, rather than a *flat structure* (Figure 1). Evidence that linguists cite in support of this hypothesis includes examples such as (1) below, in which 'one' is typically interpreted as referring to a 'red ball', rather than simply a 'ball'. Assuming anaphoric reference is limited to constituents, examples such as (1) are compatible only with the nested structure hypothesis, since only under this hypothesis does 'red ball' form a constituent. It follows that English speakers possess knowledge of the nested structure.

(1) I'll play with this <u>red ball</u> and you can play with that <u>one</u>.
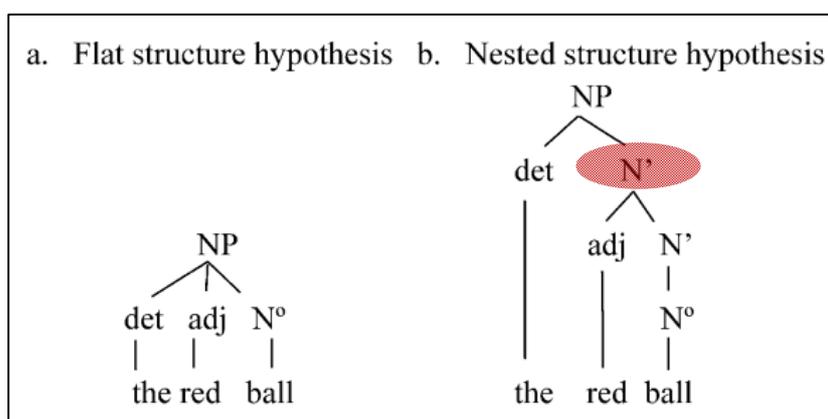

Figure 1: *Competing hypotheses for NP structure (adapted from Regier and Gahl, 2004)*

The question is how speakers of English come to possess this knowledge. LWF argue that sentences such as (1) which may be of the sort that children are exposed to, are insufficient for acquisition. Their reasoning is that any context in which (1) is felicitously uttered supports reference of 'one' to both 'ball' and 'red ball', because any suitable referent, being a red ball, is in particular a ball. Hence, any felicitous utterance would be consistent with both hypotheses, and would not allow the learner to distinguish between the two. Instead, LWF argue, unequivocal utterances such as (2) in the specified context are crucial for acquisition. The fact

that (2) can be felicitously uttered when *Max has a blue ball*, distinguishes between the two hypotheses, as it is compatible only with 'one' referring to a 'red ball', and not with 'one' referring to a 'ball', i.e. only with the nested structure hypothesis. LWF argue that such examples constitute indispensable evidence for correct acquisition.

(2) Chris has a red ball but Max doesn't have one. [*Context: Max has a blue ball*]

LWF's next step is to show that indispensable evidence such as (2) is not available to infants. In a corpus analysis comprising 55,000 parental speech utterances from the CHILDES database, LWF found that a mere 0.2% of the anaphoric usages of 'one' constituted unequivocal evidence for the nested structure. In comparison, 0.5% of the cases constituted ungrammatical usages of 'one'. LWF conclude that the indispensable evidence occurs at a level indistinguishable from noise, rendering it essentially unavailable to learners.

The final step for LWF is to argue that despite the unavailability of the critical evidence for acquisition, infants do possess the relevant knowledge (namely the correct nested structure hypothesis); they conclude that since children are not exposed to evidence which lead them to this structure, it must be the case that this syntactic knowledge is innate. To show that infants possess the relevant knowledge, they report results of the following experimental procedure.

Infants first underwent a familiarization phase in which they were presented with a single labeled item (*"Here's a yellow bottle"*). Following familiarization, two items (e.g. a yellow bottle and a red bottle) were presented to the infants who were randomly assigned a control or an anaphoric condition. In the control condition, infants heard a neutral phrase (*"Now look. What do you see now?"*) while in the anaphoric condition, subjects heard a phrase using the anaphoric expression 'one' (*"Now look. Do you see another one?"*). See Figures 2 and 3.

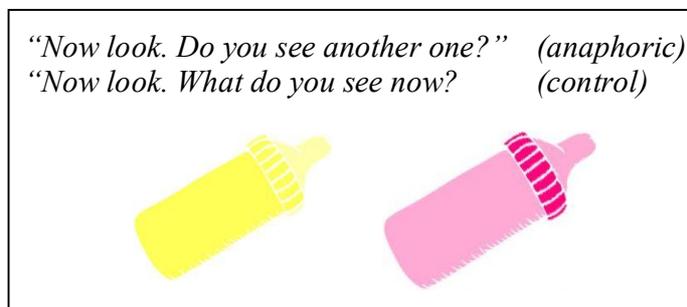Figure 2: *Familiarization phase*
(a yellow bottle is shown)

Figure 3: *Anaphoric and control conditions*
(a yellow bottle and a red bottle are shown)

Results showed that in the control condition infants preferred to look at the novel item (i.e. the red bottle in the example above) while in the anaphoric condition infants preferred to look at the familiar item (the yellow bottle). Assuming subjects prefer an item matching the linguistic stimulus (if available), LWF take this as evidence that infants know that the internal structure of NP is hierarchical, since reference to the familiar item is only possible assuming a nested structure. The significant difference between the control condition and the anaphoric condition is the reference to the yellow bottle by the NP "another one". In order for the yellow bottle to match this phrase, "one" must correspond to "yellow bottle" uttered in the familiarization

process, which must, then, form a constituent. Taken together with the unavailability of the critical data, LWF take this as support for innate syntactic knowledge.

3 Critiques, Reply and Rejoinder

Following its publication, three critiques were launched against LWF's argumentation, each challenging their work on substantially different grounds. These critiques were answered in a reply by Lidz & Waxman (2004; henceforth LW). Five years later a rejoinder was published, rekindling the debate. This section surveys the entire exchange, and concludes that all the critiques but one have been successfully rebutted.

3.1 Tomasello's Critique and LW's Reply

Tomasello (2004) challenges the syntactic nature of the findings reported by LWF, essentially arguing that the infants' performance in LWF's experiment had nothing to do with syntactic knowledge. Rather than 'one' being interpreted as an independent semantic unit in the anaphoric condition, Tomasello contends that for the infants 'another one' is a single unanalyzed unit, which the infants match against the most similar available stimulus. The meaning of 'another one', according to Tomasello, is not determined by syntax, and its interpretation can be learned from nonlinguistic contexts where it is used to refer to a similar object already under consideration. An example of the latter would be a hypothetical scenario in which a child is playing with a red block (which is not referred to linguistically) and then someone says, "Here's another one". In Tomasello's view, the expression "another one" could only refer to another red block (counter to our intuitions). Tomasello suggests that the same results would obtain if instead of hearing "Here's a yellow bottle" at familiarization, the children just heard an attention getter such as "Look at this!"

   LW's response to Tomasello is two-tiered. Firstly, they argue, only 0.37% of anaphoric uses of 'one' occur as part of the expression 'another one' with no linguistic antecedent, so such contexts are probably insufficient for acquisition. Secondly, LW point to direct empirical evidence against Tomasello's proposed account coming from Waxman and Markow (1998; henceforth WM). WM report experimental results for a procedure wherein infants were presented a yellow car with the words "See this. Look at this," (an 'attention getter' in Tomasello's terms). Crucially, the yellow car was not referred to as "a yellow car." When asked *"Can you find another one?"*, infants showed no reliable preference for the more similar yellow car over a green car. This, LWF correctly argue, is empirical evidence that in the absence of the appropriate linguistic antecedent, infants don't interpret 'another one' as referring to the most similar object, undermining Tomasello's proposal.

3.2 Akhtar, Callanan, Pullum & Scholz's Critique and LW's Reply

Akhtar, Callanan, Pullum & Scholz (2004; henceforth ACPS) challenge LWF's conclusion on two distinct grounds.

Firstly, they contest LWF's assertion that acquisition is contingent on evidence that unequivocally distinguishes between a flat structure and a nested structure (such as example (2) above). Instead, ACPS theorize that a pragmatically driven assumption could support acquisition of the nested structure hypothesis. Assume that occurrence of an adjective (e.g. 'yellow') in an antecedent NP makes it relevant for subsequent utterances by virtue of pragmatic relevance. Learners would thus assume this property ('yellow') in interpreting a subsequent use of 'one' even if it is not entailed by the sentence. As a result, even if the utterance is *compatible* with a minimal antecedent (i.e. 'one' = 'bottle') the relevance of the adjective would drive the learners to assume a more complex antecedent (i.e. 'one' = 'yellow bottle'). Additionally, ACPS argue, exophoric uses of 'one' (i.e. uses lacking a linguistic antecedent) could serve as indispensable evidence for analogous acquisition of anaphoric 'one'.

Secondly, ACPS contest LWF's unavailability argument, in that even if indispensable evidence occurs at a rate of only 0.2%, this would amount to dozens of examples across the first two years of life, which might be sufficient for learning.

Regarding acquisition being driven by pragmatic relevance, LW argue this would be much too general. If pragmatic relevance of an adjective ('yellow') in an antecedent NP makes it relevant for subsequent utterances, this should not be limited to anaphoric 'one'. Rather, such an assumption would make this property relevant also in the control condition (*"what do you see now?"*) contra LWF's findings. Also, learning by analogous exophoric evidence as suggested by ACPS, runs into the same issues that Tomasello's argument does, namely dearth of evidence and contrasting empirical findings by WM.

As for indispensable evidence accumulating over time, LW respond that along with the increasing of evidence, ungrammatical uses accumulate as well, and that since ungrammaticality isn't tagged as such, the problem remains of acquiring knowledge supported by evidence occurring at similar rates.

Some of LW's responses to ACPS seem to us to be valid. However, we do think that ACPS's point about the pragmatic relevance of the adjective guiding the child to the correct resolution of anaphoric 'one' cannot be dismissed by the assumption that it would over-generate to the control condition, as LW argue. The difference between the anaphoric condition and the control condition is precisely that the use of 'one' instructs the infants to look for something previously mentioned in the linguistic antecedent (so they prefer to look at the yellow bottle), whereas in the control condition the question presented to the infants does not (and hence, infants prefer the novel item). We will build on this insight in our account in Section 4.

3.3 Regier and Gahl's Critique and LW's Reply

Regier and Gahl (2004; henceforth RG) employ a formal statistical learning model, which they argue demonstrates that acquisition *can* be achieved given the available evidence. In a nutshell, RG argue that absence of data is itself evidence against hypotheses that predict occurrence of such data. Thus, under certain conditions, acquisition can occur even in the absence of positive evidence. To explicate, imagine a scenario in which (3) is uttered, and 'one' is understood to refer to a ball that happens to be red.

(3) I'll play with this red ball and you can play with that one.

Note, that the fact that 'one' refers to the higher N' ('red ball') is not entailed in this scenario, rather it's merely a consistent option. Other hypotheses consistent with this scenario are that 'one' refers to the lower N' ('ball') or even N^0.

Assuming that all colors are equally likely, under a lower-N' or N^0 hypothesis one would expect that over time at least in some cases the referent would turn out to be a non-red ball. However, because in reality the correct hypothesis is the higher-N' hypothesis (as demonstrated by the fact that such sentences are typically understood as referring to another red ball), non-red balls will never be attested. The more such events occur without evidence predicted by competing hypotheses (i.e. non-red balls) the less probable the competing hypotheses are. In this sense, absence of evidence predicted by competing hypotheses is negative evidence against them. Thus acquisition might occur even in the absence of unequivocal evidence.

RG present a formal statistical model that learns the correct hypothesis. Consider the probability of the event e in (4) occurring n independent times, given the hypotheses in (5).

(4) Event e = *one* refers to a red ball when *"I'll play with this red ball and you can play with that one"* is uttered.

(5) The space of hypotheses and probability of e occurring n independent times

H_1: nested structure; referent is high N' ('red ball') $\rightarrow$ $p(e^n|H_1) = \frac{1}{|H_1|^n} = \frac{1}{1^n} = 1$

H_2: nested structure; referent is low N' ('ball') $\rightarrow$ $p(e^n|H_2) = \frac{1}{|H_2|^2} = \frac{1}{\#colors^n}$

H_3: nested structure; referent is N^0 ('ball') $\rightarrow$ $p(e^n|H_3) = \frac{1}{|H_3|^2} = \frac{1}{\#colors^n}$

H_4: flat structure;   referent is N^0 ('ball') $\rightarrow$ $p(e^n|H_4) = \frac{1}{|H_4|^2} = \frac{1}{\#colors^n}$

Applying Bayes' rule, we have that $p(H|e^n) = p(e^n|H)p(H)/p(e^n)$. Assuming all hypotheses H have the same *a priori* probability, the absence of 'one' referring to non-red balls, makes $p(H_1|e^n)$ grow relative to $p(H_{2,3,4}|e^n)$ as n grows. Thus, RG conclude, the correct H_1 hypothesis might be learned even in the absence of unequivocal positive evidence.

LW's response to RG is three-tiered. First, they argue that RG's model does not fit the data children are exposed to, since in RG's model, data always includes an adjective, whereas 95% of the actual data children were exposed to have an antecedent with no adjective. Thus indirect evidence is available but will be trivialized given the overwhelming proportion of data predicted under a 'one'= N^0 hypothesis.

Secondly, LW argue that RG's model is too powerful, since it learns that 'one' must refer to the highest N', which is clearly not the case as can be seen by examples such as (6). According to RG's model, such examples are either impossible or would constitute evidence that 'one' can refer to any color, reducing the effectiveness of indirect evidence.

(6) I have a yellow bottle and you have a blue one.

Finally, LW point out that whereas LWF assumed innate aspects of syntax, RG's model assumes a very restricted hypothesis space, comprising only 4 hypotheses, so that even if RG were correct, acquisition must be extremely constrained.

3.4 Foraker, Regier, Khetarpal, Perfors & Tenenbaum's Rejoinder

Foraker, Regier, Khetarpal, Perfors & Tenenbaum (2009; henceforth FRKPT) address LW's second response to RG, namely that RG's model is too powerful in that it predicts that 'one' may refer only to the higher N' and not to any N'. FRKPT describe an alternate model, with the hypothesis space described in (7).

(7) FRKPT's hypothesis space:

H_1 'one' may refer to any N'
H_2 'one' refers to N^0

FRKPT encode these hypotheses with two grammars, identical in everything but one derivation rule for N'/ N^0, respectively, and demonstrate that a Bayesian model can learn the correct H_1. Although FRKPT's model assumes hierarchical structure, the learning they demonstrate is not trivial, since 'one'= N^0 must be excluded even if a hierarchical structure is assumed, in order to explain the complement/adjunct contrast in (8).

(8) The complement/adjunct contrast
* I'll have a piece of cheese and you can have one of apple.　　　[complement]
(OK) I want the ball with stripes and you can have the one with dots.　　[adjunct]

In sum, FRKPT demonstrate that assuming hierarchical NP structure, the correct 'one'=any N' hypothesis can be learned from available evidence.

3.5 Interim Summary

LWF claim to have shown that infants know NPs have a nested phrasal structure and that 'one' may refer to the higher N'. We believe that they have successfully rebutted Tomasello's arguments and some of ACPS's arguments. We have suggested that they have not successfully rebutted the argument that the correct resolution of 'one' can be learned by a pragmatically-driven learning strategy. As for the challenges coming from those arguing that a statistical model can learn the proper resolution of 'one', we note that the model which most successfully learns the appropriate resolution of 'one' still assumes hierarchical knowledge, and so it does not refute the existence of innate syntactic knowledge. In what follows we propose a novel critique of LWF's results. It is based on the existence of a pragmatically driven learning mechanism, as proposed by ACPS, but rather than arguing that pragmatic considerations can facilitate acquisition of the relevant syntactic knowledge, we propose that such considerations may obviate the need to assume syntactic knowledge altogether. Additionally we argue that, like in the case of FRKPT, LWF's arguments presuppose the very knowledge (hierarchy) they are trying to prove the children have. We therefore suggest that while LWF may have shown that infants know how to resolve 'one', it is unclear that this knowledge should be syntactically characterized.

4 A Novel Critique

Although 'one' is typically argued to refer to any N', in adult grammar 'one' is not entirely free to refer to any N'. Rather, by default 'one' refers to a maximal N'. To exemplify, consider the inappropriateness of a red apple being fetched after (9) is uttered.

(9) I put some green apples in the fridge. Go bring me one.

LWF's findings are consistent with this observation, since in their experiment 'one' presumably picks out the higher N'. In fact, LWF's results would be predicted even if infants' grammar allowed 'one' to refer *only* to a maximal N' (rather than to any N'). Rather than questioning the availability of sufficient evidence to support this, we ask whether in fact these results necessarily lead one to the conclusion that children possess knowledge of syntactic hierarchy. In other words, could the infants be emulating an understanding of 'one' denoting a referent specified by the higher N' without knowledge of the syntactic structure, that is, without having a notion of an N'? We argue that this in fact possible.

In their paper, WM provide evidence that infants are able to distinguish between syntactic categories, specifically, nominal and adjectival categories (N and A, respectively). Consider the possibility that for infants, anaphoric 'one' singles out a member of the category referred to by the nominal in the linguistic antecedent. Since it is a property of 'one' that it instructs the hearer to pick out a referent based on the linguistic antecedent, the infants will use whatever information present in the antecedent to determine this category. That is, the infants will take 'one' to pick out a member of the most highly specified category in the antecedent. To explicate, our proposal is given in (10).

(10) Proposal for how infants resolve the referent of 'one'
  'one' is taken to single out a member of the most specific available category

(a) When there is a linguistic antecedent consisting of a noun and an adjective, 'one' singles out a member of the category defined by A+N.
(b) When there is a linguistic antecedent with no adjective, 'one' singles out a member of the category defined by N.
(c) When there is no linguistic antecedent, a member of the category defined by a contextually recovered N is chosen, since there is no way to induce from context which adjective would be relevant for defining a more specific category.

Note that (10) is consistent not only with LWF's findings, but also with WM's at-chance results obtained in the absence of a linguistic antecedent. When presented with a yellow car and asked to find 'another one', the child can figure out the relevant nominal category, but since there are numerous possible adjectival properties which could be relevant, there is no way for the child to know what other properties may be relevant for defining the relevant category. Also note that LW's response to previous critiques is inconsequential for our proposal. Unlike Tomasello, our proposal does not involve 'another one' but rather only the resolution of 'one'. As for ACPS, our proposal is very similar to the one put forth by them, since it is also essentially pragmatic in nature. Part of learning the meaning of 'one' is learning that it instructs the hearer to rely on the linguistic antecedent in resolving its reference. Therefore, if an adjective is used in the linguistic

antecedent, it must be relevant for the resolution of 'one'. However, while ACPS argue that pragmatic relevance can constitute evidence to support learning of N', we propose that infants might not be demonstrating knowledge of NP structure altogether, but rather knowledge of how to use adjectives to calculate the category for resolution of 'one'. According to our proposal, if anything is innate it is not the knowledge of hierarchical structure but rather the notion of building the appropriate conceptual category using a noun/adjective distinction. The noun builds the basic conceptual category and the adjective builds the subcategory.

4.1 The Role of 'another'

Although not reported in the original article, LWF ran their experiment also with 'now' in the anaphoric condition instead of 'another', i.e. *"Do you see one now?"* instead of the original *"Do you see another one?"* The results obtained for this variant of the experiment are described in LW's response to Tomasello, as showing a "reliable preference" (p. 160) to the familiar object, as in the original variant (though the actual results are not reported). This means that LWF's results cannot be explained – as suggested by Tomasello – by assuming that children to do not analyze 'another one' into its constituent parts. The question is whether 'another' nevertheless has some contribution to the results. We argue that it is possible that it does, especially if we are correct in interpreting LW's phrase "a reliable preference" as indicating a less significant effect in the absence of "another". Our reasoning for this hypothetical effect-reduction is as follows.

Note that the default resolution, wherein 'one' refers to a maximal N' can be overridden, even in the presence of a linguistic antecedent, as in (11-a). However, in the presence of 'another' overriding maximality is no longer an option, (11-b).

> (11) a. I have a yellow bottle and you have a blue one.
>   b. # I have a yellow bottle and you have another blue one.
>   c. I have a yellow bottle and you have another one.

If 'another' carried no semantic import, as LWF implicitly assume, (11-a) and (11-b) should both be just as good, and (11-c) shouldn't be any better than (11-b). Instead it seems that (11-b) is degraded in much the same way as (12-a) and (12-b) are.

> (12) a. # I have a yellow bottle and you have a blue one too.
>   b. # I have a yellow bottle and you also have a blue one.

Our conclusion is that 'another' bars overriding of choosing the most specifically defined category in the presence of a linguistic antecedent. Table (13) summarizes our proposal.

(13) Proposal for how infants resolve the referent of 'one' (with and without 'another')

	- Ling. Antecedent	+ Ling. Antecedent
- another	'one' singles out a member of category N	'one' singles out a member of the most specifically defined category, unless overridden
+ another		'one' singles out a member of the most specifically defined category; no overriding

4.2 Taking Stock

Our proposal in (10) as well as its extension in (13) are consistent with 'one' referring to the N^0, obviating the need to assume knowledge of syntactic NP structure in order to account for LWF's results. As it stands, we see no reason to prefer a syntactic explanation over our explanation. In fact, the syntactic explanation for the performance of the infants in the experiments can be shown to be circular. This explanation argues that the infants have knowledge of N' because they take 'one' to refer to the string 'yellow bottle'. But notice that the conclusion that the infants have knowledge of N' rests on the assumption that anaphoric reference is possible only with constituents, and such a constraint on anaphoric resolution presupposes knowledge of hierarchy, which is exactly what we are seeking to prove. That is, the demonstration of innate syntactic knowledge depends on presupposing the existence of the very same syntactic knowledge. Barring a noncircular syntactic account, the category-based account may in fact be superior.

It is perhaps important to point out that arguments for N' in the adult grammar rest primarily on the fact that 'one' cannot refer to N^0 and must refer to a larger constituent. But the data which demonstrate this have to do with the contrast between complements and adjuncts, as in the contrast in (8) above. Such data are probably consistent *only* with a syntactic account, and not the kind of account we have put forth here. However, we know of no evidence showing that infants are sensitive to the distinction evinced by such data.

4.3 A Proposed Experiment

In order to determine whether the infants are not just calculating a category based on the intersection of properties of the N and A, but in fact do rely on some kind of structure, we propose the following experiment. As in LWF's original experiment infants would first undergo a familiarization phase in which they are presented with a single labeled item, except that now the item is modified by two distinct adjectives (*"Here's a big yellow square"*). Following familiarization, two items, each of which agree with only one of these modifiers (e.g. a big red square, and a small yellow square) are presented to the infants who are randomly assigned a

control or an anaphoric condition. In the control condition, infants hear a neutral phrase ("*Now look. What do you see now?"*) while in the anaphoric condition, subjects hear a phrase using the anaphoric expression 'one' ("*Now look. Do you see another one?"*). See figures 4 and 5. [1]
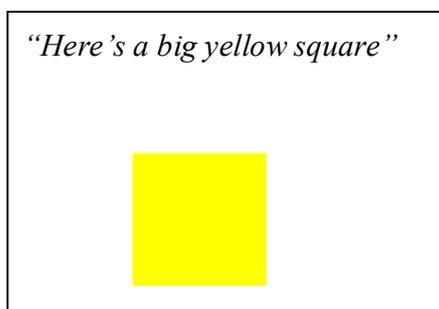

Figure 4: *Familiarization phase*
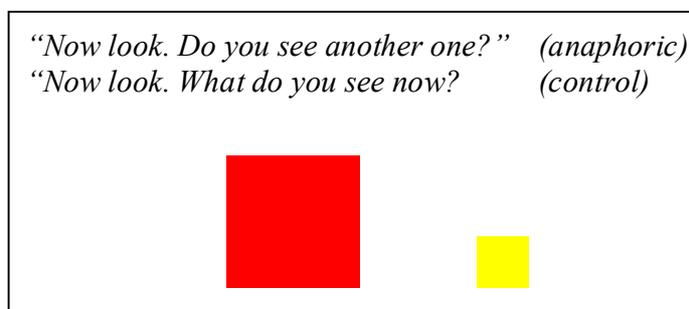*(a big yellow square is shown)*

Figure 5: *Anaphoric and control conditions*
(a big red square and a small yellow square are shown)

We would predict at-chance results for the control condition, since it doesn't contain an item identical to what was previously presented. As for the anaphoric condition, if the results turn out to indicate that infants show a preference for the item agreeing with the structurally-closer modifier (i.e. the small yellow square in figure 3) this would indicate that calculating the category based on the properties of the N and A's can't be the whole story, since in this condition, each stimulus differs from the big yellow square in one property. Rather some form of structure (even if only proximity) must be involved in the resolution of the referent. We note that our proposed experiment will still fall short of demonstrating that infants have knowledge of the correct condition on resolution, namely that 'one' can refer to any N' but not to N^0. As noted above, to demonstrate knowledge of the correct condition, one would have to show that infants are sensitive to the complement/adjunct distinction in (8).

5 Conclusion

LWF have presented experimental results showing that infants can resolve anaphoric 'one' appropriately when accompanied by 'another' (as well as 'now'). They take this to indicate that infants have knowledge of the hierarchical structure of NP as well as the fact that 'one' cannot refer to N^0.

In contrast, we have argued that another account is possible. Under the alternative account 'one' is interpreted by default to refer to a member of the most specifically defined category of the linguistic antecedent, and 'another' bans overriding of the default. In absence of a linguistic antecedent the contextually induced N defines the category. The interpretation we propose is consistent with 'one' referring to an N^0. We therefore conclude that LWF's results do not conclusively show that infants have the relevant syntactic knowledge, except perhaps for a noun/adjective distinction guiding them through a category-based resolution.

---

[1] We note that before running the proposed experiment, additional considerations must be carefully worked out; for example, it must be ensured that the competing characteristics (here, size and color) are equally salient for infants.

In sum, LWF's experiment should not be seen as supporting innate knowledge of hierarchical NP structure on its own, given the alternate category-based account presented. We proposed an experiment to tease apart categories from structure, but even such an experiment will leave open the question of whether infants can distinguish N' from N^0.

References

Akhtar, Nameera, Maureen Callanan, Geoffrey K. Pullum and Barbara C. Scholz. 2004. Learning antecedents for anaphoric one. *Cognition*, 93.

Baker, Carl Lee. 1978. *Introduction to generative-transformational syntax*. Englewood Cliffs, NJ: Prentice-Hall.

Hamburger, Henry and Stephen Crain. 1984. Acquisition of cognitive compiling. *Cognition*, 17, 85–136.

Hornstein, Norbert and David Lightfoot. 1981. Introduction. In N. Hornstein, & D. Lightfoot (Eds.), *Explanation in linguistics: The logical problem of language acquisition*. London: Longman.

Foraker, Stephani, Terry Regier, Naveen Khetarpal, Amy Francesca Perfors and Joshua B. Tenenbaum. 2009. Indirect evidence and the poverty of the stimulus: The case of anaphoric "one." *Cognitive Science* 33(2).

Lidz, Jeffrey and Sandra R. Waxman. 2004. Reaffirming the poverty of the stimulus argument: A reply to the replies. *Cognition*, 93.

Lidz, Jeffrey; Sandra R. Waxman and Jennifer Freedman. 2003. What infants know about syntax but couldn't have learned: Evidence for syntactic structure at 18-months. *Cognition*, 89.

Lightfoot, David. 1982. *The language lottery*. Cambridge, MA: MIT Press.

MacWhinney, Brian . 2000. *The CHILDES project: tools for analyzing talk (3rd ed.)*. Mahwah, NJ: Erlbaum.

Regier, Terry and Susanne Gahl. 2004. Learning the unlearnable: The role of missing evidence. *Cognition*, 93.

Tomasello, Michael. 2004. Syntax or semantics? Response to Lidz et al. *Cognition*, 93.

Waxman, Sandra R. and Dana B. Markow. 1998. Object properties and object kind: 21-month-old infants extension of novel adjectives. *Child Development*, 69.