



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Cognition 93 (2004) 157–165

COGNITION

www.elsevier.com/locate/COGNIT

Discussion

Reaffirming the poverty of the stimulus argument: a reply to the replies

Jeffrey Lidz^{a,*}, Sandra Waxman^b

^aDepartment of Linguistics, Northwestern University, 2016 Sheridan Rd., Evanston, IL 60208, USA

^bDepartment of Psychology, Northwestern University, 2029 Sheridan Rd., Evanston, IL 60208, USA

Received 6 February 2004; accepted 19 February 2004

Abstract

Lidz, Waxman, and Freedman [Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: Evidence for syntactic structure at 18-months. *Cognition*, 89, B65–B73.] argue that acquisition of the syntactic and semantic properties of anaphoric *one* in English relies on innate knowledge within the learner. Several commentaries have now been published questioning this finding. We defend the original finding by identifying both empirical and logical flaws in the critiques.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Universal grammar; Poverty of the stimulus; Language acquisition

1. Introduction

Fluent use of human language calls for a seamless integration of knowledge from a vast range of sources, both linguistic (e.g. phonology, morphology, syntax, semantics and pragmatics) and nonlinguistic (e.g. conceptual and perceptual representations of objects, events, and ideas; an appreciation of others' minds and intentions). Surely, in acquiring language, learners must coordinate information from these diverse sources. But just as surely, different aspects of human language depend on these sources to different degrees. For example, sensitivity to the intentions of others likely plays a larger role in determining the meaning of a novel word than in determining the characteristic stress pattern of an exposure language. By the same token, sensitivity to amplitude in the speech signal likely

* Corresponding author.

E-mail address: jlidz@northwestern.edu (J. Lidz).

plays a larger role in determining stress patterns than in ascribing meaning to a novel word. The trick—for the learner and for the researcher—is to consider the most appropriate sources for each of the various aspects of language.

In our original article (Lidz, Waxman, & Freedman, 2003, henceforth LWF), we focused on the acquisition of the syntactic element known as anaphoric *one*. We argued that to interpret anaphoric *one*, the learner must depend on certain aspects of phrase structure. More specifically, we provided experimental evidence that by 18 months of age, infants have knowledge of the internal phrase structure of Noun Phrases, and we provided a corpus analysis of child-directed speech to show that this knowledge could not have been gleaned from the available input. These two pieces of evidence led us to conclude that the syntactic knowledge underlying the interpretation of anaphoric *one* depends more on the child's representational presuppositions about syntax than on their experience with their language environment.

Clearly, this conclusion has struck a chord, sparking spirited responses from a strong set of commentators. These responses posed a number of interesting challenges, beckoning us to consider alternative interpretations. After a thorough consideration of each alternative, however, we find ourselves just where we began, asserting that infants at 18 months have an articulated, nested structure for the Noun Phrase and know that *one* is anaphoric to phrasal categories.

2. An overview. Poverty of the stimulus arguments

Although each commentary presents a unique perspective, they all touch upon the nature of poverty of the stimulus (henceforth POS) arguments. We therefore outline the logical form of POS arguments, place each commentary within this framework, and then respond to each in turn.

In general, the logic of the POS argument is to show that a piece of linguistic knowledge is not sufficiently triggered by the environment and hence involves some amount of innate structure in the learner. Any POS argument requires four parts. To illustrate these parts, we follow the terminology of Pullum and Scholz (2002). First, the acquirendum identifies a particular piece of syntactic knowledge. Second, the indispensability piece identifies what kind of input would be necessary for the learner to acquire the acquirendum. Third, the inaccessibility piece demonstrates that the indispensable evidence is unavailable to the learner. Fourth and finally, the acquisition piece of a POS argument shows that nonetheless the syntactic knowledge is present at the earliest possible age. Together, these pieces support the conclusion that learners have succeeded in acquiring a piece of syntactic knowledge that could not have been extracted from the environment without some inherent constraints on the hypothesis space.

Our original argument, following Baker, 1978, went like this. The acquirendum is the knowledge that *one* is anaphoric only to syntactic constituents larger than N° (i.e. the phrasal categories N' or NP). The indispensable evidence was utterances compatible only with the $one = N'$ hypothesis; utterances that are consistent with both $one = N^{\circ}$ and $one = N'$ are treated as irrelevant to learning (since they do not decide the issue). We established the inaccessibility piece of the argument by documenting in a corpus analysis that the indispensable evidence was essentially unavailable. Finally, we showed that by 18-months,

infants have indeed acquired the acquirendum; infants' behavior in our preferential looking experiment revealed that they know that *one* is anaphoric to phrasal categories.

The three responses to our original paper can be integrated within the logic of the classic POS argument. The arguments in these responses fall into two classes. One class focuses on acquisition. Engaging the work primarily on a descriptive level, Tomasello (2004) and Akhtar, Callanan, Pullum and Scholz (2004) (henceforth ACPS) ask “Do the results reveal that infants have the syntactic knowledge that LWF claim they have? Or could these results have been achieved through non-syntactic means?” A second class of challenge focuses primarily on the indispensability and inaccessibility pieces. Engaging the work on a deeper level, ACPS and Regier and Gahl, 2004 (henceforth R&G) pose questions that get at the issues of explanation and theory. Here they ask, “How did the babies develop the complex syntactic knowledge that LWF have demonstrated? Could this syntactic knowledge have been learned from the input?” We first address the descriptive challenge, defending our finding and clarifying its force. This clears the deck for a fuller treatment of the rich theoretical issues that lie at the heart of the POS argument.

3. Descriptive challenges. Have we really documented syntactic knowledge in infants?

Both Tomasello and ACPS consider whether infants' performance in our experiments really depends upon syntactic knowledge. ACPS raise this concern within the context of a nuanced discussion that also involved issues of explanation and theory (see below). But Tomasello, who focuses exclusively on this issue, asserts that infants' performance on our task reveals nothing in the way of syntactic knowledge on the part of our infant subjects.

Tomasello's argument is based on the supposition that *another one* is an undecomposable expression (i.e. it is an unanalyzed whole) that is “conventionally associated” with “certain types of nonlinguistic situations” and that it therefore has a meaning that is not determined by its syntax. More precisely, he asserts that the expression *another one* simply refers to any object that is similar to an object already under consideration. To illustrate, Tomasello offers 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. He then asserts that scenarios like these provide a sufficient basis for acquiring the meaning of *another one*. Our objection takes three parts.

First, although the expression *another one* is not uncommon in the input to children (in the corpus used in LWF, it accounted for 33 (or roughly 4%) of the 792 anaphoric uses of *one*), only rarely is it used without a linguistic antecedent (only 3 of these 33 cases had no antecedent, or 0.37% of the anaphoric uses). This observation blunts Tomasello's intuition that such contexts are sufficient to the acquisition of *another one*.

Second, our intuitions regarding such hypothetical scenarios does not match Tomasello's. In our view, in such scenarios, the expression “Here's another one”, could just as easily refer to a blue block, or to another toy altogether, depending upon the objects available and the goals at hand.

Third, and perhaps most telling, Tomasello's intuition is contradicted directly by empirical evidence. Waxman and Markow (1998) presented this very scenario to

21-month-old infants in a control condition for a word-learning experiment. An experimenter introduced control infants to a target object (e.g. a yellow car) while saying, “See this. Look at this.” The experimenter then presented two test objects (e.g. a different yellow car and a green car) and asked, “Can you find *another one*?” As in Tomasello’s scenario, there was no linguistic antecedent for *one*. If Tomasello was correct, then infants in this condition should pick the test object that most closely resembled the target object; in this case, they should reveal a clear preference for the yellow car. But this was not the case. Infants had no such preference. In fact, they chose the yellow car at a rate that was indistinguishable from chance. This stands as empirical evidence that infants do not interpret *another one* as referring to the most similar object available in the context.

In contrast to these infants in Waxman and Markow’s control condition, infants in the LWF experiment did hear a linguistic antecedent during familiarization and did pick the most similar object at test. This indicates that LWF’s infants attended to the structure of the NP presented during familiarization, and that their appreciation of the internal syntactic structure of that NP led them (but not Waxman and Markow’s control infants) to pick the most similar object at test.

One further piece of evidence is relevant here. Although we did not include this evidence in the original report for lack of space, we did run a control condition in which we left out *another* in the test question. As in the original study, infants were familiarized to an object (e.g. a yellow bottle) which was labeled (here’s a yellow bottle...). However at test, these infants heard “do you see one now?” If Tomasello’s intuition is right, then infants in this condition should show no preference, since on this view the preference in our original experiment was due to the noncompositional meaning of the expression *another one*. Following Tomasello’s logic, *one* by itself is not conventionally associated with the kinds of nonlinguistic experiences that *another one* is and so does not uniquely identify either of the bottles. But, as in the original experiment, infants showed a reliable preference for the yellow bottle. This suggests that performance in our original experiment cannot be explained by appealing to the “kinds of nonlinguistic experiences conventionally associated with the expression *another one*.” Rather, infants’ interpretation of *one* reflects their knowledge that this expression is anaphoric to a phrasal category.

4. Theoretic and explanatory challenges. How do infants acquire the syntax of anaphoric *one*?

The other two commentaries were concerned more with the question of how children manage to acquire the syntactic knowledge that supports the appropriate interpretation of anaphoric *one*. At issue in these commentaries is whether we successfully established the indispensability and inaccessibility portions of our argument.

4.1. Akhtar, Callanan, Pullum and Scholz (ACPS)

ACPS question our conclusions from multiple perspectives, challenging both the indispensability and inaccessibility portions of our POS argument. And within the context of these questions, they challenge the acquisition piece of the argument, asking whether

we have in fact demonstrated that infants have syntactic knowledge of the structure of NP or the properties of anaphoric *one*.

4.1.1. *The indispensability argument*

We argued that the only evidence that would lead the learner to the correct hypothesis regarding anaphoric *one* (nested structure in NP, *one* = N') would be events in which *one* is compatible with this hypothesis alone. ACPS take a broader view of the input, arguing that there are two additional kinds of indispensable evidence (namely, syntactically ambiguous utterances and exophoric reference) that could support learning of anaphoric *one*. Their argument is predicated on the view that comprehension is driven by the relevance principle (Sperber & Wilson, 1995).

ACPS argue that “if learners do sometimes have to figure out from primary data that *one* has a multi-word antecedent, it does not follow that the context must *entail* that the antecedent is multi-word. Mere pragmatic implication or conversational implicature will do just as well.” In other words, by virtue of the principle of relevance, the use of *yellow* in an antecedent noun phrase makes the property *yellow* relevant to all subsequent utterances. It therefore follows that learners assume this property (even in situations which do not entail that it is included in the antecedent) and interpret *one* accordingly. In essence, then, the argument is that a pragmatically driven assumption can then support the nested structure grammar in which *one* is anaphoric to N'.

But this argument runs aground because the assumption in this relevance-oriented view is too broad. The assumption is that the adjective in the antecedent should make the property relevant not only in cases of anaphoric *one* (as in our experimental condition: *do you see another one?*), but also in a broad range of other subsequent utterances¹ (as in our control condition: *what do you see now?*). For both types of utterances, a relevance-based approach predicts (a) that the color of the object presented during familiarization was relevant (otherwise, why would we have mentioned it) and (b) that participants would look to the yellow bottle in both conditions. But the data from LWF do not support this prediction. Infants in the experimental condition preferred the yellow bottle at test, but those in the control condition performed at chance. This difference between the experimental and the control conditions demonstrates that infants' interpretation of *one* cannot be reduced to the principle of relevance.

ACPS also argue that exophoric uses of *one* (i.e. uses in which the referent of *one* is inferred from context alone) could serve as indispensable evidence to the acquisition of anaphoric *one*. They state, “if *Hand me that* is correctly understood (to mean the yellow bottle) and the next utterance is *give me another one*, no syntactic constituent serves as antecedent at all. But the relevance principle still suggests that *one* has, say, *yellow bottle* as antecedent rather than *just bottle*...Such uses may be highly instructive to the learner.”

This challenge parallels the concern raised by Tomasello. To revisit that issue briefly, recall that the evidence from Waxman and Markow (1998) showed that exophoric uses are

¹ Unless, of course, the subsequent utterance contradicts the first utterance, as in “You have a yellow bottle and I have a blue one.”

not informative in this situation. This renders implausible the assertion that exophoric uses of *one* form the basis for learning the anaphoric uses.

4.1.2. *The inaccessibility argument*

In LWF, we used an analysis of a CHILDES corpus to address this issue of inaccessibility. Our analysis revealed that technically speaking, the input does contain some situations which could, in principle, support the acquisition of anaphoric *one*. The trouble is that such utterances occurred at rate of 0.2%, which is the very same rate as ungrammatical uses of *one*. The crux of our argument was this: Certainly learners hear examples that they could learn from, but they hear these at the same rate as ungrammatical examples that they shouldn't learn from.

ACPS take a different direction, arguing that the 0.2% of situations that would allow learning would amount to dozens of examples accumulated over the first 2 years of life. We concur, but our claim was not about raw percentages or even projections from the corpus to the actual input a child might receive over a more extended period of time. For if the number of good examples increases *n*-fold over a period of time, so does the number of ungrammatical uses. Unfortunately, the input to the learner does not come marked as to whether and how it should be used by the learning algorithm. The raw input is not marked with regard to grammaticality, felicity, truth value, etc. By logic, then, we can assume that if learners do not acquire the syntactic properties of anaphoric *one* from the 0.2% of the data that is ungrammatical (of which we are certain, since mature speakers of English clearly have command of the syntax), then neither could they acquire the syntactic properties from the 0.2% of the data that is crucially informative.

4.2. *Regier and Gahl (R&G)*

R&G's criticism takes yet a different tack, focusing on the indispensability (and consequently, inaccessibility) parts of our argument. They correctly point out an important component of any learnability problem, namely, that indirect negative evidence can have a powerful influence on acquisition (Chomsky, 1981). R&G say,

Given evidence that is consistent with several hypotheses, a learner can come to discriminate among them, for principled, domain-general reasons. In particular, if one of the hypotheses predicts not only the input that is seen, but also input of another sort that is never seen, that *absence* can serve as evidence against the hypothesis.

R&G's point is that learners may be able to use the absence of certain kinds of data as evidence against hypotheses that would have predicted such data to occur. We agree, in principle. The trouble is that the solution proposed by R&G fails. To show why this is so, we outline three problems inherent in their proposal which, singly and in combination, lead us back to our original point—that the acquisition of anaphoric *one* cannot be learned from the available input (even if the learner considers the full range of negative and positive evidence).

First, the data presented to the R&G model bears little relation to the input seen by actual children. Their model saw only examples of *one* being used with an antecedent containing

an adjective (e.g. here's a red ball. do you see another one?). Their model learned quickly that *one* = [_N red ball]. But this not altogether surprising since there were no exposures in which the actual referent was not red. Child learners are not so fortunate. In the corpus we examined, examples like those presented to the R&G simulation (with *one* being used with an antecedent adjective) were rare. Less than 5% of the utterances containing *one* were preceded by antecedents containing adjectives; 95% of the data they were exposed to had an antecedent that did not contain an adjective. That is, the vast majority of cases were more like "here's a bottle. do you see another one?" This raises questions concerning the degree to which the R&G model fits the child's case in acquiring anaphoric *one*.

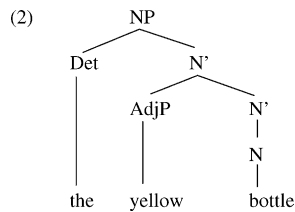
Second, even if we accept R&G's model and apply it to the data that children actually do receive, we suspect that the simulation will run aground. This is because in cases with no adjective in the antecedent, there is no a priori reason for the learner to privilege any one particular property of the referent (e.g. color) over another. That is, in cases like "here's a bottle" (where the bottle happens to be yellow), there is no reason to interpret "get another one" (anaphoric *one*) as referring to a yellow bottle, because that referent (the yellow bottle) is also a full bottle, a warm bottle, your sister's bottle, etc. In fact, learners should conclude that *one* in such cases makes no commitment to any of a host of potential properties of the referent. Put simply, in the face of input that more closely matches that provided to actual children, *one* would most plausibly be interpreted as referring to the object category (bottle), making no commitment to its various properties. And because utterances of this type are so frequent in the actual input to infants, this will have a consequence for learner's interpretation when the antecedent NP does mention color (e.g. here's a yellow bottle), size, shape, texture, attractiveness, temperature, or any other of its properties. Now, if for some reason, the learner were paying attention to the various properties of the bottle in the cases where it was not mentioned, she might decide that properties of the antecedent (like color) are altogether irrelevant for assigning an antecedent to *one*, thereby swamping the potential effect of indirect negative evidence in the cases when the adjective is present and therefore potentially relevant. Our point here is not that indirect negative evidence is unavailable to a learner. We believe that it is. Rather, we claim that the indirect negative evidence that kids really do get could not solve the learning problem in this instance, given the overwhelming proportion of data which are consistent with the *one* = N° hypothesis and the vast range of properties that should correctly be ignored almost always when assigning an antecedent to anaphoric *one*.

A third issue brought up by R&G's model concerns the details of the grammar of anaphoric *one*. Our original discussion was primarily concerned with the *possibility* of *one* taking a multi-word antecedent, e.g. *yellow bottle* in our example. It is this possibility that provides the evidence against a flat structure grammar. R&G's discussion, however, assumes that it is *necessary* for *one* to have a multi-word antecedent if one is available. This is transparently not the case, as shown by examples like (1a) and (1b).

- (1) a. I have a yellow bottle and you have a blue one.
b. I like the yellow bottle but you like that one.

Now, these cases do involve anaphora to N', but they differ from our original example in that here *one* is anaphoric to the lower of the two N's, i.e. the one containing only

the string *bottle*. Consider the structure in (2).



Here, there are two nodes labeled N' and either one of these is a potential antecedent for anaphoric *one*.² The fact that *one* can refer to either of two N' nodes is important because it shows that R&G's learning model is too powerful. The conclusion of their model is that *one* must refer to the largest available N' , namely the one containing the string *yellow bottle*. But this conclusion is too strong, because *one* can refer to lower instances of N' , as we have seen. It is simply not an entailment of sentences like "I like this red ball and you like that one" that *one* be anaphoric to *red ball*. Although this reading may be the most felicitous *one* in some contexts, in other contexts and utterances, *one* can indeed refer to the lower N' . If we accept R&G's indirect evidence model, then such contexts either would be ruled impossible (which would be an error in learning) or would constitute evidence that *one* can refer to balls of any color, even when the antecedent mentions color, and would thus blunt the effectiveness of any indirect negative evidence of the sort described by R&G. In short, the use of indirect negative evidence in the fashion proposed by R&G

² We know that the nonbranching N' dominating *bottle* is present because of the impossibility of *one* being anaphoric to complement taking nouns:

- (i)
 - a. I met the member of congress...
 - b. *...and you met the one of the rotary club
 - c. [NP the [N' [N member] [PP of congress]]]
- (ii)
 - a. I reached the conclusion that syntax is innate...
 - b. *...and you reached the one that learning is powerful
 - c. [NP the [N' [N conclusion] [CP that syntax is innate]]]

These contrast with cases in which what follows the head noun is an adjunct/modifier. Here, *one* can substitute for what appears to be only the head noun.

- (iii)
 - a. I met the student from Peoria...
 - b. ...and you met the one from Podunk.
 - c. [NP the [N' [N' [N student]] [PP from Peoria]]]
- (iv)
 - a. I met the student that you invited to the party
 - b. ...and you met the one that Bill invited
 - c. [NP the [N' [N' [N student]] [CP that you invited to the party]]]

What these cases tell us is that *one* can take a single word as its antecedent only when that single word is dominated by a nonbranching N' . In other words, in cases like (1a), it must be the case that *ball* = N' , as in the structure in (2). If it weren't, we would have no way to distinguish this case from one in which *one* cannot substitute for a single word, as in (i) and (ii).

leads to a grammar in which it is never possible for *one* to refer to the lower of two N's, contrary to fact. Their model is simply too restrictive.

A fourth, more general, issue brought up by R&G's reply concerns the constraints on the model. We argued that two aspects of syntactic structure guide the acquisition process: (a) that NPs have a nested (X-bar theoretic) structure, and (b) that anaphora to X° be excluded from the hypothesis space. R&G argue that these particular pieces of linguistic knowledge need not be innate. However, R&G do have a massively restricted hypothesis space (one containing only four hypotheses). So, we are in agreement that the hypothesis space for syntactic acquisition must be severely restricted. We differ only in which hypotheses are represented in that space. If it turns out, as R&G suggest, that the learner can work through that hypothesis space using domain general learning mechanisms, that would be a very interesting finding, but it is important to emphasize, as R&G do themselves, that all learning must be constrained. Without a severely restricted hypothesis space, the learner cannot even get started.

5. Conclusion

We are pleased that our research has sparked such enthusiastic discussion from such a diverse source of commentators. The commentaries have forced us delve deeper into the nature of the POS and to consider alternatives to our particular instantiation of the POS argument. But, having visited these alternatives, we end right where we began. By the time they are 18-months old, infants have an articulated, nested, structure for the Noun Phrase and they know that *one* can be anaphoric only to phrasal categories. As far as we can tell, this grammatical knowledge could not have come from the input that they were exposed to (even if they use pragmatic information and/or negative evidence to evaluate that input). The alternative, then, is that this knowledge must have come from within the infants themselves.

References

- Akhtar, N., Callanan, M., Pullum, G., & Scholz, B. (2004). Learning antecedents for anaphoric *one*. *Cognition*, 93, 141–145.
- Baker, C. L. (1978). *Introduction to generative-transformational syntax*. Englewood Cliffs, NJ: Prentice-Hall.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: Evidence for syntactic structure at 18-months. *Cognition*, 89, B65–B73.
- Pullum, G., & Scholz, B. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, 19, 9–50.
- Regier, T., & Gahl, S. (2004). Learning the unlearnable: The role of missing evidence. *Cognition*, 93, 147–155.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed). Cambridge, MA: Blackwell.
- Tomasello, M. (2004). Syntax or semantics? Response to Lidz et al. *Cognition*, 93, 139–140.
- Waxman, S., & Markow, D. (1998). Object properties and object kind: 21-month-old infants extension of novel adjectives. *Child Development*, 69, 1313–1329.