

Why language acquisition is a snap

STEPHEN CRAIN AND PAUL PIETROSKI

Abstract

Nativists inspired by Chomsky are apt to provide arguments with the following general form: languages exhibit interesting generalizations that are not suggested by casual (or even intensive) examination of what people actually say; correspondingly, adults (i.e., just about anyone above the age of four) know much more about language than they could plausibly have learned on the basis of their experience; so absent an alternative account of the relevant generalizations and speakers' (tacit) knowledge of them, one should conclude that there are substantive "universal" principles of human grammar and, as a result of human biology, children can only acquire languages that conform to these principles. According to Pullum and Scholz, linguists need not suppose that children are innately endowed with "specific contingent facts about natural languages." But Pullum and Scholz don't consider the kinds of facts that really impress nativists. Nor do they offer any plausible acquisition scenarios that would culminate in the acquisition of languages that exhibit the kinds of rich and interrelated generalizations that are exhibited by natural languages. As we stress, good poverty-of-stimulus arguments are based on specific principles – confirmed by drawing on (negative and crosslinguistic) data unavailable to children – that help explain a range of independently established linguistic phenomena. If subsequent psycholinguistic experiments show that very young children already know such principles, that strengthens the case for nativism; and if further investigation shows that children sometimes "try out" constructions that are unattested in the local language, but only if such constructions are attested in other human languages, then the case for nativism is made stronger still. We illustrate these points by considering an apparently disparate – but upon closer inspection, interestingly related – cluster of phenomena involving: negative polarity items, the interpretation of 'or', binding theory, and displays of Romance and Germanic constructions in Child-English.

1. Introduction

Before getting down to brass tacks, let us sketch two perspectives on the topic of this special issue. The first is a Chomskian view, which we endorse. On this view, human languages exhibit interesting and unexpected generalizations. The linguist's job is to find them and provide theories that explain why these generalizations hold. However, the utterances speakers make, along with the conversational contexts in which they make them, do not reveal the theoretically interesting linguistic phenomena – much less the deeper principles that help unify and account for these phenomena. Moreover, facts about expressions that speakers *don't* use and the meanings that speakers' *don't* assign to (otherwise well-formed) expressions, are at least as important as facts about what speakers actually say, and the meanings they actually assign. As generalizations over this large initial data set emerge, across various languages, linguists propose and test hypothesized grammatical principles that would (in the course of normal human experience) give rise to languages that exhibit the phenomena being described. But inquiry is hard. Many of the relevant facts appear to be contingencies of human psychology, which may well have been shaped in part by demands imposed on it by the kinds of signals and interpretations that human minds can process; and little is known about these demands on human psychology, especially on the meaning side of the equation. Moreover, the space of logically possible grammatical principles is immense. The principles discovered so far describe a tiny fraction of the space of possible languages, despite having impressive empirical coverage with regard to actual human languages. So theorists search for constraints that would allow for *only* the relatively small number of languages that humans can naturally acquire and use.

Fortunately, linguists can draw upon a vast amount of positive and negative evidence – both within and across languages – and they can consider all of the evidence at once. Nevertheless, there remains a familiar tension between explanatory power and descriptive adequacy: good theories do not merely summarize observed facts; and in a complex world, any nontrivial generalization introduces an “idealization gap” between theory and data. As in all areas of scientific enquiry, considerations of simplicity and theoretical economy are relevant in linguistic theories. In addition, linguistic theories must be compatible with observations about the nature of children's cognitive abilities, as well as their histories of linguistic stimulation. For all normal children acquire adult linguistic competence despite the considerable latitude in environmental input to different children. So any principles posited as descriptions of languages spoken by adults must be such that they are acquirable by *any* normal child who undergoes the kind of experience that the corresponding adults underwent. Any principles that are posited must also be learnable in the time

frame that characterizes normal language acquisition. Regardless of how hard it is for linguists to discern grammatical principles, research in child language has shown that young children know them, often by age three, and sometimes younger than that. Yet children of any age are rarely if ever exposed to negative data; they do not have access to cross-linguistic facts; and they presumably do not (as linguists often do) confirm hypothesized principles based on how well these principles unify and explain disparate phenomena – such as island effects, weak-crossover effects, etc. So either children can (in a way that linguists cannot) extract grammatical principles from what adults actually say, in the circumstances in which they say them, or children do not have to learn such principles. In so far as the former option seems implausible, one is led to the conclusion that children know basic grammatical principles largely by virtue of their innate biological endowment. This suggests that the linguist's task is to characterize (i) the initial state – often called “Universal Grammar” – that children bring to the task of language acquisition, and (ii) the possible modifications of that initial state, under the influence of experience, into specific grammars of the sort that are exhibited by mature language users.¹

In the target article, Pullum and Scholz suggest another view of the relation between linguistic theory and the primary linguistic data available to children, as well as a different view of the cognitive apparatus children bring to the task of language learning. In their view, the grammatical principles of natural languages are not much deeper than is suggested by the evidence available to children. Indeed, the principles underlying natural languages are simple enough that children – and even some linguists – can figure them out. From this perspective, one “would question whether children learn what transformational-generative syntacticians think they learn.” Since the principles underlying natural languages don't run deep, children do not have to be super scientists to learn them; the evidence available is adequate to the inferential task children face. Pullum and Scholz offer a few illustrations of the “shallow” records that children could keep of their linguistic experience. These are piecemeal records of construction types. These construction types are learned solely from positive evidence, using an intuitively simple grammatical typology. Such record keeping obviates the need for children to learn linguistic principles by “intensely searching” for evidence or by considering vast arrays of both positive and negative evidence, within and across languages, simultaneously. In short, piecemeal acquisition of construction types avoids the kind of “over-hasty” generalizations that would require “unlearning” on the basis of negative evidence.

1. We use ‘grammar’ to talk about psychological properties of speakers. If one finds this terminology objectionable, one can substitute Chomsky's (1981) term ‘I-grammar’ (internal grammar).

Of course, advocates of nativism agree that children don't learn grammatical principles in the way that linguists discern them. The nativist claim is that while a child's experience figures in the explanation of why she acquires the local language(s), as opposed to others, the grammatical principles that describe the space of possible human languages – and thus constrain the child's path of language acquisition – are not learned at all. Trivially, positive evidence suffices for language *acquisition*, given the innate cognitive apparatus children bring to the task. On the alternative view, positive evidence suffices for language *learning*. According to Pullum and Scholz, linguists need not suppose that children are innately endowed with “specific contingent facts about natural languages.” If the data available to children are rich enough for them to determine the grammatical rules of natural languages, given the right inferential techniques, then appeals to other sources of data (i.e., innately specified principles) are at best a useful crutch for theorists – and at worst a source of erroneous claims about alleged “gaps” between the facts concerning particular languages and the evidence available to children.

We think this second perspective faces a cluster of theoretical and empirical problems. For starters, in order to provide a genuine alternative to the Chomskian view, one has to say a lot more than Pullum and Scholz do about the relevant inferential techniques. Some have suggested that familiar methods of “data-sifting”, either by traditional induction or more intricate methods of statistical sampling, would lead any normal child to the state of grammatical knowledge achieved by adult speakers; where the data is described without recourse to a sophisticated linguistic theory (e.g., Tomasello 2000). But in our view, this proposal has never been confirmed for any of the domains that are good candidates for innate linguistic knowledge (e.g., syntax and semantics). To their credit, Pullum and Scholz avoid empiricism of the most implausible sort. They talk more about “construction types,” thus granting that children impose at least some grammatical categories on their experience. But this raises the question of why children impose *certain* construction types as opposed to others. It is also worth asking why natural languages, and hence children, manifest *certain* generalizations as opposed to others, and why the generalizations that do turn up govern disparate phenomena, and reach across different linguistic communities. If the proposal is that children know which construction types to use when constructing records of adult speech, for purposes of figuring out how the local language works, this isn't something nativists need to deny – at least not if the space of possible construction types turns out to be immense, as compared with the range of construction types actually exploited by children.

But whether the focus is on construction types, or grammatical principles that constrain the available constructions in natural languages, the empirical question is whether the proposed linguistic principles are fruitful in explaining the range of facts that natural languages exhibit, and that children acquire

rapidly in the absence of much experience. Unfortunately, Pullum and Scholz don't consider examples that illustrate the tight relation between the details of linguistic theory and the most impressive poverty-of-stimulus arguments. It is no accident, in our view, that the most impressive poverty-of-stimulus arguments present specific analyses of several linguistic phenomena, following a general discussion of how knowledge of language goes beyond experience; see, e.g., Chomsky 1981, 1986; Hornstein and Lightfoot 1981. For the best nativist arguments are based on independently confirmed claims about adult grammars: evidence suggests that adults know some abstract generalization, G – concerning binding, head-movement, or whatever; one asks whether children could have plausibly learned G on the basis of evidence available to them; if that seems unlikely, and experimental investigations reveal that very young children know G, one tentatively concludes that G is due, at least in large part, to Universal Grammar. (See Crain and Pietroski 2001 for a recent review.)

In attempting to rebut poverty-of-stimulus arguments, Pullum and Scholz are free to wonder whether young children really know the grammatical principles in question. But a *response* to nativists requires an alternative account for a growing body of psycholinguistic evidence. It also calls for either (a) an alternative linguistic theory according to which *adults* don't know the relevant generalizations, or (b) a remotely plausible account of how children could acquire the adult knowledge on the basis of evidence available to them. We're not saying that this *can't* be done; but, we are saying that poverty-of-stimulus arguments are not rebutted until this is done. And we see no way for children to learn, in ways that Pullum and Scholz gesture at, the interesting and unexpected generalizations that linguists working in the transformational-generative tradition have discovered.² We present some of these generalizations in Sections 2 and 3 below.

Another problem for the second perspective is that it cannot explain the patterns of non-adult constructions that crop up in the speech of young children. We take this up in Section 4. While no one thinks that children advance immediately to a grammar that generates and interprets constructions in the same way as adults in the same linguistic community, the two views lead to quite distinct expectations about the “deviant” constructions that children are likely

2. Perhaps we are not criticizing the rebuttal by Pullum and Scholz of what they take to be the argument from the poverty-of-stimulus (APS). But if so, then so much the worse for them in focusing on that particular version of the APS. No charitable reader of Chomsky could think that *his* arguments for nativism are supposed to be independent of the detailed grammatical theories that he has defended. Indeed, we find it hard to see how one could advance an interesting version of linguistic nativism that is independent of specific claims about the grammatical knowledge of adults: until one has a firm grip on what adult linguistic competence is like, one can't even begin to hypothesize about the cognitive equipment that children would need (in addition to their experience) to achieve adult-like states.

to exhibit. On the first view, there are natural seams (or parameters) of natural language, and child speech should follow these seams, even when it diverges from the speech of adults. Children will, under the pressure of experience, explore some part of the space of humanly possible languages; but they will never “try out” a language that violates core principles of Universal Grammar. By contrast, given what Pullum and Scholz say, the obvious prediction is that children’s constructions should simply be less articulated ones than those of adults. Children should initially try out “simple” construction types that may need to be refined in light of experience. As we’ll see in Section 3 below, the evidence from studies of child language favor the nativist view, and resist explanation on the view taken by Pullum and Scholz.

2. Empirical details

We now turn to some empirical details. These illustrate the problems that beset the perspective we attribute to Pullum and Scholz. There are many phenomena we could discuss in this context. But since we want to display the form of (what we take to be) a good poverty-of-stimulus argument, we will focus in some detail on just one cluster of closely related facts.

Let’s start with some much-discussed facts concerning negative polarity items (NPIs) like *any*, *ever*, or the idiomatic *a red cent*. The appearance of such items is perfectly fine in many linguistic contexts, but somehow wrong in others. The following ten examples illustrate a small fraction of the construction types that permit negative polarity items: sentences with negation (1) or negative adverbs (2); prepositional phrases headed by *before* (3) or *without* (4); antecedents of conditionals (5); verb-phrases headed by *forbid* (6) or *doubt* (7); the first argument of *no* (8) and its second argument (9); the first argument of *every* (10). The oddity of example (11) illustrates that NPIs are not licensed in the second argument of *every*.³

- (1) *I don’t talk to any other linguists.*
- (2) *I never talk to any other linguists.*
- (3) *I usually arrive at the gym before any other linguist wakes up.*
- (4) *I went to the gym without any money.*
- (5) *If any linguist goes to the gym, I go swimming.*

3. One can specify the meaning of a quantificational expression using (something like) set-theoretic relations. On this view, a quantificational expression in a simple declarative sentence names a relation between two sets: first, there is the set picked out by the NP (e.g., ‘linguist with any brains’ in (8)); second, there is the set picked out by the VP (e.g., ‘admires Chomsky’ in (8)). We will refer to these as the first and second arguments of the quantifier.

- (6) *I forbid any linguists to go swimming.*
 (7) *I doubt that any linguist can refute Chomsky.*
 (8) *No linguist with any brains admires Chomsky.*
 (9) *No linguist has any brains.*
 (10) *Every linguist with any brains admires Chomsky.*
 (11) **Every linguist has any brains.*⁴

Perhaps the kind of piecemeal acquisition advocated by Pullum and Scholz could let children learn all the positive environments in which NPIs can appear, e.g., (1)–(10). On this account, moreover, children could avoid introducing NPIs in the second argument of *every*, as in (11), simply because they do not encounter such sentences. One is still left to wonder *why* only the first argument of *every* licenses NPIs, and not its second argument, and how children successfully discern such apparently subtle distinctions amidst the buzzing bloom of conversation. Still, if children are meticulous and conservative record-keepers who encounter (all of) the relevant examples, they would not need further information about which constructions do not license negative polarity items. But a little more investigation reveals that there is much more to explain than just the distributional facts about where NPIs are actually licensed.

Children must also learn how to interpret disjunctive statements. In the vast majority of cases, English sentences with the disjunction operator *or* are naturally understood with an “exclusive-*or*” interpretation – implying that *not both* disjuncts are satisfied (see, e.g., Braine and Romain 1981, 1983, who claim that only the exclusive-*or* reading is available for adults, as well as children).

- (12) *You may have cake or ice cream for dessert.*
 (13) *Eat your veggies or you won't get any dessert.*

4. We restrict attention, in the present discussion, to *any* on its “true universal” as opposed to “free choice” uses (see, e.g., Horn 2000; Kadmon and Landman 1993; Ladusaw 1996). While speakers may assign an interpretation to *I went to lunch with any money* – i.e., I went to lunch with any money at hand – the use of *any* in that construction clearly contrasts with that in *I went to lunch without any money*. Some relevant contrasts to (1–10) include the following degraded constructions (setting aside free-choice uses): *I talk to any other linguists*; *I usually arrive after any other linguist wakes up*; *if I go swimming, any other linguist goes to the gym*; *some linguist with any brains admires Chomsky*; *some linguist admires any philosopher*. Using other negative polarity items, compare *I never paid a red cent for that book* and *every linguist who ever disagreed with me likes you* with the degraded *I paid a red cent for that book* and *some linguist who ever disagreed with me likes you*. Finally, while *I think any linguist can refute Chomsky* sounds fine (arguably because it involves the use of free-choice *any*), compare *I doubt you ever paid a red cent for that book* with the terrible *I think you ever paid a red cent for that book*.

However, there are also disjunctive construction types that can only be understood with a “conjunctive” interpretation. For example, a disjunctive construction with negation, such as *not (A or B)*, is understood to be equivalent in meaning to *(not A) and (not B)*. Despite the abundance of exclusive-*or* interpretations of disjunctive statements in the input to children, the exclusive-*or* reading of the disjunction operator cannot be the source of the conjunctive interpretation of disjunctive statements, because the negation of a disjunctive statement using exclusive-*or* is true if both disjuncts are satisfied. In forming the conjunctive interpretation of disjunctive statements (e.g., with negation), children must somehow ignore the available evidence from “positive” statements with disjunction, in which exclusive-*or* is favored, as in (12) and (13); children must somehow learn to use inclusive-*or* instead, at least when interpreting disjunctive statements with negation.⁵

How do children navigate through their linguistic experience to discover when to assign an exclusive interpretation to disjunctive statements, and when not to? To answer this, it pays to look at a list of construction types that exhibit the conjunctive interpretation of disjunction. Here are ten of the relevant constructions: sentences with negation (14) or negative adverbs (15); prepositional phrases headed by *before* (16) or *without* (17); antecedents of conditionals (18); verb-phrases headed by *forbid* (19) or *doubt* (20); the first and second arguments of *no* (21) and (22); and the first argument of *every* (23). Example (24) illustrates that the second argument of *every* permits an exclusive-*or* interpretation of disjunction, so this linguistic environment does not require a conjunctive interpretation.

- (14) *I don't talk to linguists or philosophers.*
- (15) *I never talk to linguists or philosophers.*
- (16) *I try to get to the gym before the linguists or philosophers.*
- (17) *I go to the gym without the linguists or philosophers.*
- (18) *If a linguist or a philosopher goes to the gym, I go swimming.*
- (19) *I forbid linguists or philosophers from going to the gym.*
- (20) *I doubt the linguists or the philosophers can refute Chomsky.*
- (21) *No linguist or philosopher admires Chomsky.*
- (22) *No one with any brains admires linguists or philosophers.*

5. Our use of the term ‘reading’ is not intended to commit us to the view that *or* is ambiguous in English, or that disjunction is ambiguous in any natural language. As we discuss shortly, it is reasonable to suppose that the meaning of *or* conforms to that of disjunction in standard logic (i.e., inclusive-*or*), but that statements with *or* are often judged to be true only in a subset of its truth conditions, namely those that are associated with exclusive-*or*. Similar remarks apply to the meaning of *any* (see Footnote 4).

- (23) *Every linguist or philosopher with any brains admires Chomsky.*
 (24) *Everyone admires a linguist or a philosopher.*

Clearly, there is considerable, perhaps complete overlap between the constructions in which NPIs appear, and those in which disjunctive statements receive a conjunctive interpretation – and cannot be interpreted using an exclusive-*or* reading of disjunction. This is, presumably, not a coincidence. To discover the generalization, however, linguists needed to amass a large amount of data – and then relate facts about licit grammatical forms (involving NPIs) to facts about possible interpretations (of disjunctive statements) (Chierchia 2000). Unsurprisingly, making the generalization more precise requires a more intensive scrutiny of data, both positive and negative.

But even if we suppose the overlap between licensing of NPIs and interpretation of disjunction is exhaustive, there would still be several ways to state the generalization. One “shallow” formulation would be a description of the evidence, as in (25).

- (25) The conjunctive interpretation is assigned to disjunctive statements if and only if an NPI can appear in that linguistic environment.

An alternative formulation of the generalization ties it to a (deeper) property of the relevant environment, that of *downward entailment*. Ordinary declarative sentences license inferences from subsets to sets, as in (26a). This is called an ‘upward’ entailment. By contrast, a downward entailing linguistic environment licenses inferences from sets to their subsets (Ladusaw 1996). For example, sentential negation creates a downward entailing linguistic environment, as illustrated by the obviously valid inference in (26b).

- (26) a. *Noam bought an Italian car.* \Rightarrow *Noam bought a car.*
 b. *Noam didn't buy a car.* \Rightarrow *Noam didn't buy an Italian car.*

Notice that the first, but not the second, argument of *every* is a downward entailing environment: If every linguist bought a car, then it follows that every Italian linguist bought a car, but it doesn't follow that every linguist bought an Italian car.

Using the same diagnostic for downward entailment, we find that *no* is downward entailing in both of its argument positions; that *some* is downward entailing in neither of its argument positions; and so on. We are now positioned to move beyond the descriptive generalization in (25), and in the direction of an explanation, along the lines of (27).

- (27) Downward entailing linguistic environments license NPIs *and* constrain the interpretation of disjunctive statements (to conjunctive readings) (cf. Horn 1989: 234).

Much work remains. One wants to know *why* downward entailment constrains both NPI licensing and the interpretation of disjunctive statements. And why *these* constraints? Ludlow (2002) explores with ingenuity the suggestion that NPIs are, as their name suggests, indeed licensed by the presence of negation – and that despite surface appearances, all of the licensing environments involve an element of negation at the level of Logical Form (cf. Laka 1990). Chierchia (2000) proposes that the so-called exclusive readings of disjunctive statements in examples like (12) and (13) result from a kind of Gricean implicature that is computed within the human language system. The idea is that a sentence with a scalar term has both a “basic” meaning and a “derived” meaning, where the derived meaning is determined by conjoining the basic meaning with the negation of a corresponding statement in which the basic scalar operator is replaced with the next strongest operator on the scale. If the derived meaning is more informative than the basic meaning, then a speaker using the sentence will be heard as “implicating” the more informative claim. For example, the logical operators ‘ \vee ’ (inclusive disjunction) and ‘ $\&$ ’ (conjunction) form a simple scale; the latter is stronger than the former, since ‘ $A \& B$ ’ is true only if ‘ $A \vee B$ ’ is true, but not vice versa. If ‘ \vee ’ gives the basic meaning of *or*, the derived meaning of disjunctive statements of the form *A or B* is given by ‘ $(A \vee B) \& \text{not } (A \& B)$ ’ – which is equivalent to ‘ A exclusive-or B ’. On this view, *or* always stands for inclusive disjunction, but the derived meaning of *A or B* is more informative than its basic meaning; correspondingly, a speaker who says *A or B* will be heard as making a claim with the following implication: *not (A and B)*. However, if *or* appears in the scope of negation, as in (21), the derived meaning is not more informative than the basic meaning.⁶

Both the proposal by Ludlow and the one by Chierchia strike us as plausible. Regardless of whether either of them is correct, however, we see no reason to doubt the truth of what all the research in this area suggests: that some semantic principle – call it ‘downward entailment’ – unifies what otherwise seem to be the disparate phenomena of NPI licensing, the interpretation of disjunctive statements, and the validity of inferences like the one in (26b).

One could deny all this, of course, and reject the claim that human grammars (which children attain) are properly characterized by any deep generalization like (27). Perhaps only the descriptive generalization in (25) – or no generalization at all – is correct. But simply denying apparent generalizations is just bad science. One can’t avoid nativist conclusions by refusing to do linguistics. And we don’t think Pullum and Scholz would advocate this approach. Instead, we suspect that they would offer an alternative proposal about what children

6. In natural languages, *not (A or B)* is equivalent to *not (A) and not (B)* as in de Morgan’s laws; but the derived meaning of *not (A or B)* – ‘ $\text{not } (A \vee B) \& \text{not } [\text{not } (A \& B)]$ ’ – would be a contradiction, and thus not a viable interpretation of *not (A or B)*.

know when they know the descriptive generalization in (25). But any such principle that is empirically equivalent to (27) will provide the basis for a poverty-of-stimulus argument, absent a credible account of how all normal children could learn the principle. Perhaps one can avoid direct appeal to (Universal Grammar) constraints concerning downward entailment by saying that children record what they hear in terms of abstract construction types that respect (25). But then the question reduces to why children deploy *those* construction types, as opposed to others.

The challenge for Pullum and Scholz, therefore, is to describe a plausible acquisition scenario (e.g., for the descriptive generalization in (25)), according to which children avoid uncorrectable overgeneralizations, without supposing that children approach the acquisition process with specific linguistic knowledge of the sort they regard as unnecessary (e.g., the linguistic property of downward entailment). So far as we can tell, Pullum and Scholz offer no hint of how to formulate a learning account that eventuates in attainment of the specific linguistic knowledge that nativists tend to focus on, such as downward entailment. In short, it's not enough to mention ways in which children could learn some things without Universal Grammar. To rebut poverty-of-stimulus arguments, one has to show how children could learn what adults actually know; and as close investigation reveals, adults know a lot more than casual inspection suggests. *That* is the nativist's main point.

3. Further empirical details

Let's continue in this vein a bit further. Even given a characterization, say in terms of downward entailment, of which construction types license NPIs, further work remains. For example, we have seen that certain constructions with a negative element, such as *not*, license NPIs, such as *any* (see (1) above). But one wants to know *how* the negative element needs to be related to the NPI in order to license it. One logical possibility is that the NPI *any* is licensed in constructions in which *not* precedes *any*. But in both of the following examples, *not* precedes *any*, whereas *any* is licensed only in the second example.

(28) *The news that Noam had not won was a surprise to some/*any of the linguists.*

(29) *The news that Noam had won was not a surprise to some/any of the linguists.*

Other (deeper) generalizations have been proposed to explain the licensing of NPIs in constructions with negation. One proposal is stated in (30) (see Fromkin et al. 2000: Chapter 4).

(30) Negation must c-command an NPI to license it.

C-command is an abstract structural relationship that cannot be defined in terms of perceptible features of word strings.⁷ One can try to formulate a more shallow generalization that could be learned, not based on c-command. One possibility, similar in kind to analyses that Pullum and Scholz seem to endorse, would be something along the lines of (31), where (31a) illustrates a construction type in which *some*, but not *any*, is permitted; by contrast, (31b) is a construction type in which both *some* and *any* are permitted.

- (31) a. ... neg+V+V+NP+P+some
 b. ... V+neg+NP+P+some/any

Of course, one is left to wonder how children know to keep records of this sort, as opposed to others. It seems implausible, to say the least, that children are recording everything they hear and searching for every possible pattern. Do children *learn* to apply category labels like NP, V and P, or is this part of the cognitive apparatus human beings are disposed to project onto their experience?

But even setting such questions aside, the proposal that c-command is the relevant structural relationship for the licensing of NPIs has much to recommend it, as opposed to the construction type approach advocated by Pullum and Scholz. The c-command account has unexpected and independent support from a host of other linguistic constructions. Consider (32), for example.

- (32) a. *The bear who laughed never expected to find any dogs at the party.*
 b. **The bear who never laughed expected to find any dogs at the party.*

In (32a), the negative adverb *never* c-commands *any*, but not in (32b). Correspondingly, only (32a) is acceptable. Adopting the Pullum and Scholz approach, one could suppose that children encode the facts in (32) in terms of construction types, where another construction type that permits NPIs is one of the form *never+V+INF+...*, but NPIs would not be encountered in constructions of the form *never+V+V+INF+...* But even if some such proposal could describe the facts, record keeping of this kind fails to explain *why* NPIs are licensed in the first type of construction, but not in the second; and it fails

7. While some linguists seem to use the licensing of NPIs as a diagnostic of c-command, its precise definition and the level of representation at which it applies (d-structure, s-structure, LF, semantic representation) is the subject of considerable debate (see, e.g., the papers in Horn and Kato 2000).

to tie this fact together with the fact that NPIs are licensed in (31b), but not in (31a).

There are ample reasons for thinking that c-command plays a crucial role in the interpretation of these constructions, and in many other constructions where the licensing conditions for NPIs is not at issue (see Epstein et al. 1998). To take a familiar kind of example, the pronoun *he* cannot be referentially dependent on the referring expression, *the Ninja Turtle*, in (33); whereas this relationship is possible in (34); and referential dependence is only possible, in (35), between the reflexive pronoun *himself* and the referring expression, *the father of the Ninja Turtle* (but not *Grover* or *the Ninja Turtle*).

- (33) *He said the Ninja Turtle has the best smile.*
- (34) *As he was leaving, the Ninja Turtle smiled.*
- (35) *Grover said the father of the Ninja Turtle fed himself.*

One standard explanation for the prohibition against referential dependence in (33) is that a pronoun cannot be referentially dependent on a referring expression that it c-commands. In (34), the pronoun does not c-command *the Ninja Turtle*, so anaphoric relations are permitted. In addition, reflexive pronouns must be referentially dependent on a ‘local’ antecedent that c-commands it, as (35) illustrates.

The account in which c-command governs the interpretation of pronouns has been extended to pronouns in Wh-questions, as in the so-called strong crossover question in (36). The wh-question in (36) is unambiguous; the pronoun must be interpreted deictically, as indicated by the boldface in (36a). Notice that, in addition to the deictic reading of the pronoun, another reading is possible in (37), according to which the pronoun *he* is “bound;” on this reading (indicated in 37b by the underlining) the question asks: for which *x*, did *x* say that *x* has the best smile?

- (36) *Who did he say has the best smile?*
 - a. **Who** did **he** say has the best smile
 - b. *Who did he say has the best smile
- (37) *Who said he has the best smile?*
 - a. **Who** said **he** has the best smile (Deictic reading)
 - b. Who said he has the best smile (Bound Pronoun reading)
- (38) *He said the Ninja Turtle has the best smile.*
- (39) *He said who has the best smile*
- (40) *Who did he say t has the best smile*

An account invoking c-command has been proposed to explain the absence of the bound pronoun reading of strong crossover questions like (36). Chomsky (1981) proposed that the strong crossover Wh-question in (36) is derived from an underlying representation that mirrors the declarative sentence in (38), as indicated in (39). In the course of the derivation, the Wh-word *who* “crosses over” the pronoun on its way to its surface position, leaving behind a “trace” in its original position, as in (40). The illicit reading would be possible only if (a) the pronoun could be referentially dependent on the trace or (b) the trace could be referentially dependent on *he* (while *he* was also somehow bound by *who* – perhaps by virtue of its link to the trace). Chomsky proposed that traces of Wh-movement cannot be referentially dependent on expressions that c-command them; in this (limited) sense, traces are like names and definite descriptions. Again, even if this specific proposal is not correct, it seems hard to deny that there are generalizations – at quite a distance from the data – that need to be stated in terms of abstract notions like c-command. This invites the question of how children learn that (36) is unambiguous whereas (37) is ambiguous. Empirical evidence that children know the relevant linguistic principles by age three is presented in Crain and Thornton (1998), and there is a growing body of experimental evidence showing that young children obey the licensing conditions on the negative polarity item *any*, and compute the conjunctive interpretation of disjunction in downward entailing linguistic environments (Crain, Gualmini and Meroni 2000; Musolino, Crain and Thornton 2000; Thornton 1995).

Pullum and Scholz owe a plausible acquisition scenario of the same phenomena, which ties together the licensing of NPIs and the interpretation of pronouns in declaratives and in Wh-questions. From their perspective, the learning account should not be based on the assumption that children start the acquisition process with specific linguistic knowledge of the sort ascribed by nativists. Assuming that Pullum and Scholz would not choose to deny that languages harbor interesting and unexpected generalizations, they owe a plausible linguistic theory that captures (and at least starts to explain) these generalizations without the apparatus that generative-transformational grammarians posit, which include the notion of c-command. We repeat: there is a tight relation between the *details* of workaday linguistics, the *specific* principles that seem to govern human languages, and the best poverty-of-stimulus arguments. Because Pullum and Scholz avoid the details of linguistic theory and their role in explaining linguistic generalizations, their rebuttal to poverty-of-the stimulus arguments misses the mark.

4. Patterns of non-adult constructions

Thus far, we have been pressing (what we take to be) familiar kinds of nativist considerations. A less obvious problem for the kind of learning scenario advanced by Pullum and Scholz concerns the pattern of non-adult constructions that appear in the language of young children. Other things being equal, Pullum and Scholz should predict that children (in so far as they diverge from adults) will initially employ constructions that are less articulated than those employed by adults. Complexity in the child's hypotheses about the local language should be driven by what the child hears; otherwise, complex hypotheses will look like reflections of a mental system that imposes certain structures more or less independently of experience.

But according to the perspective of linguists working within the generative-transformational tradition, children should be expected to sometimes follow developmental paths to the adult grammar that would be very surprising from a data-driven perspective. Of course, any normal child quickly internalizes a grammar equivalent to those of adults around them. But a child who has not yet achieved (say) a dialect of American English can still be speaking a natural language – albeit one that is (metaphorically) a foreign language, at least somewhat, from an adult perspective. And interestingly, the children of English speakers often do exhibit constructions that are not available in English – but ones that are available in other languages spoken by actual adults. This is unsurprising if children project beyond their experience, rather than being inductively driven by it. From a nativist perspective, children are free to try out various linguistic options (compatible with Universal Grammar) before ‘setting parameters’ in a way that specifies some particular natural grammar, like that of Japanese or American English. A natural extension of this line of thought is sometimes called the *Continuity Hypothesis* (Pinker 1984; Crain 1991; Crain and Pietroski 2001). According to the Continuity Hypothesis, child language can differ from the local adult language only in ways that adult languages can differ from each other. The idea is that at any given time, children are speaking a possible (though perhaps underspecified) human language – just not the particular language spoken around them. If this is correct, we should not be surprised if children of monolingual Americans exhibit some constructions characteristic of German, Romance or East Asian languages, even in the absence of any evidence for these properties in the primary linguistic data. Indeed, such mismatches between child and adult language may be the strongest argument for Universal Grammar.

We conclude with two examples: Wh-questions that reveal a trace of Romance in Child-English; and Wh-questions that reveal a trace of Germanic in Child-English. In each case, the relevant facts come into view only when they are framed within a detailed theory of some non-English phenomena, along-

side some otherwise puzzling observations about Child-English. We think these phenomena constitute real poverty-of-stimulus arguments, but they are arguments that rest on details of the sort that Pullum and Scholz don't consider.

To set the stage for the account of Child-English that attributes to it properties of Romance, we briefly review Rizzi's (1997) analysis of matrix and long-distance Wh-questions in Italian. On Rizzi's analysis, wh-questions are formed by I-to-C movement, in which an inflectional node (I) moves to a commanding position that is associated with complementizers (C). The obligatory nature of such movement is illustrated by the fact that certain adverbs (e.g., *già*, *ancora*, and *solo*) cannot intervene between the wh-operator and the inflected verb in interrogatives:

- (41) *Che cosa hanno già fatto?*
 What have-3PL already done
 'What have (they) already done?'
 (42) **Che cosa già hanno fatto?*
 what already have-3PL done?
 'What already have (they) done?'

By contrast, the intervention of these adverbs is tolerated in declaratives, because no movement has taken place.

- (43) *I tuoi amici hanno già fatto il lavoro.*
 the-PL your friends have-3PL already done the-SG work
 'Your friends have already done the work.'
 (44) *I tuoi amici già hanno fatto il lavoro.*
 the-PL your friends already have-3PL done the-SG work
 'Your friends already have done the work.'

In contrast to ordinary wh-elements, the wh-words *perché* (why) and *come mai* (how come) do not require I-to-C movement in matrix questions. Example (45) shows that the adverb *già* as well as the entire subject NP may intervene between the question-word and the inflected verb.

- (45) *Perché (I tuoi amici) già hanno finito*
 Why (the-PL your friends) already have-3PL finished
il lavoro?
 the-SG work
 'Why (your friends) already have finished the work?'

To explain the contrast, Rizzi proposes that *perché* and *come mai* are not moved at all; rather, they are base generated in a position that is intrinsically endowed with the syntactic property that makes movement necessary for ordinary wh-elements.

However, in deriving long-distance Wh-questions with *perché* and *come mai*, these Wh-elements must move; hence, they should behave like ordinary wh-elements. That is, even *perché* and *come mai* are expected to block the intervention of short adverbs or a subject NP in long-distance questions. The evidence for I-to-C movement in long-distance is subtle; it involves the interpretation of questions, rather than their form. Consider the examples in (46) and (47). Example (46) is ambiguous. On one reading, *perché* is locally construed; this reading asks about the reason for the event of “saying”. But, in addition, (46) has a long-distance reading, which is about the reason for the resignation. In example (47), by contrast, there is only one reading, on which *perché* is construed locally. A long-distance reading is unavailable because a subject intervenes between the wh-element and the inflected verb, revealing that the wh-element *perché* could not have moved from the embedded clause.

- (46) *Perché ha detto che si dimetterà?*
 why have-3SG said that self resign-3SG.future
 ‘Why did he say that he would resign?’
- (47) *Perché Gianni ha detto che si dimetterà*
 why Gianni have-3SG said that self resign-3SG.future
(non a Piero)?
 (not to Piero)
 ‘Why did Gianni say that he will resign (not to Piero)?’

This brings us, at last, to Child-English. It has frequently been noted that *why*-questions in Child-English tend to lack (subject-auxiliary) inversion to a greater extent than other wh-elements, and that the absence of inversion for *why*-questions persists in children’s speech well after inversion is consistently present in other wh-questions. Adopting the Continuity Hypothesis, de Villiers (1990) and Thornton (2001) have both suggested that children of English-speaking adults initially treat the question-word *why* in the same way as Italian adults treat *perché* or *come mai*. That is, children of English-speaking adults base generate *why* in a structural position that differs from the position occupied by other wh-expressions. On this view, children base generate *why* in a position that does not require I-to-C movement – unlike other wh-elements. This explains the absence of inversion in Child-English.

Following the Rizzi-style analysis, Child-English should nevertheless require inversion for long-distance *why*-questions, even if a particular child does not require inversion for matrix *why*-questions. If this is correct, such a child should differ from English-speaking adults in the way he forms matrix *why*-questions (without inversion), but the child should parallel English-speaking adults in producing well-formed long-distance *why*-questions. From a data-driven perspective, this pattern of (non)conformity is surely not anticipated.

Precisely this pattern was found, however, in an experimental and longitudinal diary study by Thornton (2001), who recorded questions by one child, AL, between the ages of 1;10 and 4;6. By age 3;4, AL required inversion for all matrix wh-questions except for *why*-questions; non-inversion persisted in (over 80% of) AL's matrix *why*-questions for more than a year after that, as illustrated in (48).

- (48) *Why you have your vest on?*
Why she's the one who can hold it?
Why it's his favorite time of day?

By contrast, from the time AL was 3-years-old until she reached her fourth birthday, she produced 65 long-distance *why*-questions, and only seven of them lacked inversion. The remaining 58 were adult-like, as were all 39 long-distance questions with wh-elements other than *why*.

- (49) *Why do you think you like Cat in the Hat books?*
Why do you think mummy would not wanna watch the show?
What do you think is under your chair?
How do you think he can save his wife and her at the same time?

But, as we saw (example (47)), I-to-C movement is not required in Italian, as long as the question receives a local construal, rather than a long-distance reading. Two of AL's seven *why*-questions without inversion are presented in (50); it seems likely in both cases that AL intended the local construal of *why*.

- (50) *Why you think Boomer's cute? I'm cute too.*
Why you said there's no trunk in this car?

In short, the production data suggest that an English-speaking child can analyze *why*-questions like the corresponding questions are analyzed in Romance languages, such as *perché* in Italian. In producing matrix *why*-questions, however, AL was ignoring abundant evidence in the input indicating a mismatch between her grammar and that of adults in the same linguistic community. Nevertheless, AL did not violate any principles of Universal Grammar. In particular, AL adhered to the grammatical principles that require I-to-C movement in all long-distance questions.

Another example of children's non-adult (but UG-compatible) productions is the "medial-wh" phenomenon, which reveals a trace of Germanic in Child-English. Using an elicited production task, Thornton (1990) found that about one-third of the 3–4 year-old English-speaking children she studied consistently inserted an 'extra' wh-word in their long-distance questions, as illustrated in (51) and (52) (also see Crain and Thornton 1998; Thornton 1996, for a description of the experimental technique used to elicit long-distance wh-questions, from children as young as 2;7).

- (51) *What do you think what pigs eat?*
 (52) *Who did he say who is in the box?*

This “error” by English-speaking children is presumably not a response to the children’s environment, since medial-wh constructions are not part of the primary linguistic data for children in English-speaking environments. However, structures like (51) and (52) are attested in dialects of German, as the example in (53) illustrates (from McDaniel 1986).

- (53) *Wer_i glaubst du wer_i nach Hause geht?*
 who-NOM think-2SG you who-NOM towards house go-3SG
 ‘Who do you think who goes home?’

Further investigation shows that the similarity of Child-English to a foreign language runs deep. For both adult Germans and American children, lexical (full) wh-phrases cannot be repeated in the medial position. For example, German-speaking adults judge (54) unacceptable, and English-speaking children never produced strings like (55), as indicated by the ‘#.’ Instead, children shortened the wh-phrase or omitted it altogether, as in (56) (Thornton 1990).

- (54) **Wessen Buch_i glaubst du wessen Buch_i Hans liest?*
 who-GEN book think-2SG you who-GEN book Hans read-3SG
 ‘Whose book do you think whose book Hans is reading?’
 (55) *#Which Smurf do you think which Smurf is wearing roller skates?*
 (56) *Which Smurf do you think (who) is wearing roller skates?*

Finally, children never used a medial-wh when extracting from infinitival clauses, so they never asked questions like (57). Nor is this permissible in languages that allow the medial-wh.

- (57) *#Who do you want who to win?*

This complex pattern of linguistic behavior suggests, once again, that many children of English-speakers go through a stage at which they speak a language that is like (adult) English in many respects, but one that is also like German in allowing for the medial-wh. There is nothing wrong with such a language – it just happens that adults in the local community do not speak it.

To conclude, the non-adult linguistic behavior of children is relevant for the recent debate on nativism. The evidence we reviewed from Child-English suggests that children may often fail to match their hypotheses to the input in ways that run counter to the kind of account of language acquisition offered by Pullum and Scholz. Instead, children appear to be free to project unattested

hypotheses, so long as incorrect hypotheses can later be retracted, presumably on the basis of positive evidence. If the ways in which child and adult language can differ is limited to ways in which adult languages can differ from each other, then this would be compelling evidence in favor of the theory of Universal Grammar. On the account we envision, children's linguistic experience drives children through an innately specified space of grammars, until they hit upon one that is sufficiently like those of adult speakers around them, with the result that further data no longer prompts further language change.

University of Maryland

References

- Braine, Martin D. S. and Barbara Romain (1981). Development of comprehension of 'or': evidence for a sequence of competencies. *Journal of Experimental Child Psychology* 31: 46–70.
- (1983). Logical reasoning. In *Handbook of Child Psychology, vol. 3: Cognitive Development*, John Flavell and Ellen Markman (eds.), 46–70. New York: Academic Press.
- Chierchia, Gennaro (2000). Scalar implicatures and polarity phenomena. Paper presented at NELS 31, Georgetown University, Washington, DC.
- Chierchia, Gennaro, Stephen Crain, Maria Teresa Guasti, and Rosalind Thornton (1998). "Some" and "or": a study on the emergence of logical form. In *Proceedings of the Boston University Conference on Language Development 22*, Annabel Greenhill, Mary Hughes, Heather Littlefield, and Hugh Walsh (eds.), 97–108. Sommerville, MA: Cascadilla Press.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.
- Crain, Stephen (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences* 14: 97–650.
- Crain, Stephen and Paul Pietroski (2001). Nature, nurture and Universal Grammar. *Linguistics and Philosophy* 24 (2): 139–186.
- Crain, Stephen and Rosalind Thornton (1998). *Investigations in Universal Grammar. A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, MA: The MIT Press.
- Crain, Stephen, Andrea Gualmini, and Luisa Meroni (2000). The acquisition of logical words. *LOGOS and Language* 1: 49–59.
- deVilliers, Jill (1990). Why questions? In *Papers in the Acquisition of Wh: Proceedings of the UMass Roundtable*, Thomas L. Maxfield and Bernadette Plunkett (eds.), 155–171. Amherst, MA: University of Massachusetts Occasional Papers.
- Epstein, Samuel, Groat Erich M., Kawashima Ruriko, and Kitahawa Hisatsugu (1998). *A Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.
- Fromkin, Victoria (ed.) (2000). *Linguistics: An Introduction to Linguistic Theory*. Malden, MA: Blackwell Publishers.
- Grice, H. Paul (1975). Logic and conversation. In *Syntax and Semantics 3: Speech Acts*, Peter Cole and James Morgan (eds.), 41–58. New York: Academic Press.
- Heim, Irene (1984). A note on negative polarity and downward entailingness. *Proceedings of the North East Linguistic Society*, 14: 98–107.
- Horn, Laurence (1989). *A Natural History of Negation*. Chicago, IL: University of Chicago Press.
- (2000). Pick a theory (not just any theory). In *Negation and Polarity: Syntactic and Semantic Perspectives*, Laurence Horn and Yasuhiko Kato (eds.), 147–192. Oxford: Oxford University Press.

- Horn, Laurence and Yasuhiko Kato (eds.) (2000). *Negation and Polarity: Syntactic and Semantic Perspectives*. Oxford: Oxford University Press.
- Hornstein, Norbert and David Lightfoot (1981). Introduction. In *Explanations in Linguistics: The Logical Problem of Language Acquisition*, Norbert Hornstein and David Lightfoot (eds.), 9–31. London: Longman.
- Kadmon, N. and F. Landman (1993). Any. *Linguistics and Philosophy* 16: 353–422.
- Ladusaw, William (1996). Negation and polarity items. In *Handbook of Contemporary Semantic Theory*, Shalom Lappin (ed.), 321–342. Oxford: Blackwell.
- Laka, Itziar (1990). Negation in syntax: on the nature of functional categories and projections. Unpublished Ph.D. dissertation, MIT, Cambridge.
- Lightfoot, David W. (1991). *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: MIT Press.
- Ludlow, Peter (2002). LF and natural logic: the syntax of directional entailing environments. In *Logical Form and Language*, Gerhard Preyer and George Peter (eds), 132–168. Oxford: Oxford University Press.
- McDaniel, Dana (1986). Conditions on wh-chains. Ph.D. dissertation, City University of New York.
- May, Robert (1985). *Logical Form*. Cambridge MA: MIT Press.
- Munn, Alan B. (1993). Topics in the syntax and semantics of coordinate structure. Unpublished Ph.D. dissertation, University of Maryland, College Park.
- Musolino, Julien, Stephen Crain, and Rosalind Thornton (2000). Navigating negative semantic space. *Linguistics* 38: 1–32.
- Pinker, Steven (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.
- Progovac, Ljiljana (1994). *Negative and Positive Polarity: A Binding Approach*. Cambridge, MA: Cambridge University Press.
- Rizzi, Luigi (1997). The fine structure of the left periphery. In *Elements of Grammar: Handbook of Generative Syntax*, Liliane Haegeman (ed.), 281–337. Dordrecht: Kluwer Academic Publishers.
- Thornton, Rosalind (1990). Adventures in long-distance moving: the acquisition of complex wh-questions. Ph.D. dissertation, University of Connecticut, Connecticut.
- (1995). Children’s negative questions: a production/comprehension asymmetry. In *Proceedings of ESCOL*, J. Fuller, H. Han, and D. Parkinson (eds.), 306–317. Ithaca, NY: Cornell University.
- (1996). Elicited production. In *Methods for Assessing Children’s Syntax*, Dana McDaniel, Cecile McKee, and Helen S. Cairns (eds.), 77–102. Cambridge, MA: MIT Press.
- (2001). Two tasks. Paper presented at the 2nd Annual Tokyo Conference of Psycholinguistics, Workshop, Keio University, Tokyo.
- Tomasello, Michael (2000). Do young children have adult syntactic competence? *Cognition* 74: 209–253.