JOHN COLLINS

COWIE ON THE POVERTY OF STIMULUS

ABSTRACT. My paper defends the use of the poverty of stimulus argument (POSA) for linguistic nativism against Cowie's (1999) counter-claim that it leaves empiricism untouched. I first present the linguistic POSA as arising from a reflection on the generality of the child's initial state in comparison with the specific complexity of its final state. I then show that Cowie misconstrues the POSA as a direct argument about the character of the pld. In this light, I first argue that the data Cowie marshals about the pld does not begin to suggest that the POSA is unsound. Second, through a discussion of the so-called 'auxiliary inversion rule', I show, by way of diagnosis, that Cowie misunderstands both the methodology of current linguistics and the complexity of the data it is obliged to explain.

1. INTRODUCTION

About half of Fiona Cowie's *What's Within* (1999) seeks to undermine Chomsky's long argued for hypothesis that the human comprehension and use of language is subserved (in combination with other cognitive systems) by a *language faculty*, viz., a dedicated component of the mind/brain that represents a rich and complex system of innate information that severely constrains the nature and development of possible human languages.[1] For Cowie (2001, 239), this picture of linguistic competence has "acquired the status, almost, of dogma"; the task of *What's Within* is not to show that Chomsky is wrong, still less is it to argue for behaviourism or any other species of empiricism, rather, its brief is to dispute the grounds of the 'dogma': everything, more or less, is still to play for.

Cowie's (1999) has received some strong criticism (e.g., Fodor, 2001; Matthews, 2001).[2] This assault, however, has focused on questioning the logical landscape Cowie depicts. Suffice it to say that I share many of the criticisms offered by Fodor and Matthews. Even so, such a focus lets Cowie off a certain hook. Her main criticism, as I see it, is directed at the so-called *poverty of stimulus* argument (POSA). This argument is non-demonstrative, empirical, and Cowie challenges it on precisely such grounds. Cowie (2001, 244–245) complains (against Matthews) that, without an empirical or methodological counter challenge, her argument stands, for it is Chomsky and others who are basing bold claims on POS

considerations, not her. This riposte is somewhat myopic. Matthews corrects Cowie on her understanding of formal learnability results, which she concedes; he also raises sound objections to Cowie's discussion of 'negative evidence', to which she neglects to respond. In the sequel, I shall press home this latter offensive and show (i) that the data Cowie marshals does not begin to suggest there is anything wrong with the POSA as commonly employed and (ii) that Cowie misconstrues the methodology of the generative program to such an extent that it becomes quite trivial that the POSA is unsound; once read aright, however, the program will be seen to be immune to Cowie's challenge.

2. NATIVISM: PRELIMINARIES

As indicated, Cowie is no empiricist; her claim is just that various forms of empiricism are not in fact refuted by the standard nativist arguments (essentially, POSA). An *enlightened empiricism*, for example, may postulate a faculty that trades in representations and is dedicated in its contents and operations to the domain of language. This is what makes the position enlightened, what makes it empiricist is that it claims that the principles which go to determine the correct output of the faculty are learnt (or, perhaps better, acquired) as opposed to encoded innately. Cowie's own tentative position is *weak nativism*, a thesis that accepts representationism, domain specificity, and nativism, but rejects UG (Cowie appears to confuse UG – the initial state of a component of the mind/brain – with whatever Chomsky's latest theory of the component is). Cowie's contention is that much of the perceived force of the arguments and data for Chomsky's claims derive from a too simple taxonomy; it is because, in the extreme case, a rejection of UG is associated with a crude behaviourism or empiricism that the thesis seems so unassailable, for to reject UG, under such an association, entails a rejection of the much weaker claims of innateness, domain specificity, and representationism. Likewise, a rejection of innateness is taken to be concomitant with a rejection of domain specificity and representationism. Going the other way, Cowie's claim is that an argument which refutes behaviourism is not ipso facto one which refutes an enlightened empiricism, still less a weak nativism. Indeed, as we shall see, Cowie argues that even empiricism *simpliciter* (a position which rejects domain specificity but holds to representationism and innate general principles of learning) is immune to POSA. Cowie's argument, then, is not that empiricism is true, nor does she take herself to have demonstrated the falsity of the Chomskyan position. Her conclusion is merely that, as far as the POSA (inter alia) demonstrates, empiricism is still a live option.

There is a dialectical difficulty here. So modest is Cowie's general conclusion that, in a certain sense, it would be churlish to dispute. After all, the questions are empirical ones, and given the relatively immature state of linguistics and cognitive science in general, both theoretically and evidentially, any form of dogmatism is not apt, and no-one should gainsay future research. Such innocent concessions, however, are quite uninteresting: everything we presently believe about linguistic cognition might be false, but we are not therefore rationally free to follow any avenue we please without some form of corroboration that the path leads to the truth. This line has been forcefully followed by Fodor (2001) and Matthews (2001). Whatever the dialectical situation might be, however, Cowie's claim still apparently stands that the POSA for the Chomskyan position is really quite poor, so poor that it even admits empiricism. Let us first look at the POSA, then consider Cowie's rejoinder.

3. POVERTY OF STIMULUS

The POSA for linguistic nativism is directly motivated by perhaps the most general observation we can make about humans and language: barring congenital defect or later trauma, we all end up using (at least) a *particular* language, although we *might* have ended up using *any other* language if our formative linguistic stimuli would have been other than they were. In other words, our brains, in distinction to those of rabbits or chimpanzees, are such as to enable us to acquire language *as such*, although they are not primed to acquire any particular language: take an English neonate, move it 24 miles across the Channel and, *ceteris paribus*, it will end up speaking French, not English. Consequently, whatever cognitive equipment we begin with – whatever distinguishes us from rabbits and chimpanzees in the relevant respect – must be specific enough (*not* necessarily specific to language) for us to arrive at English or French or Walpiri, etc., while also being general enough to target any language with equal ease. On reflection, this is obvious enough, and does not itself entail any particular story about linguistic cognition apart from the bland remark that humans begin with innate equipment, whether specific to language or not, that enables them to acquire any language. I say this is bland, because it only really rules out the absurd idea that all constraints on language acquisition are exogenous, the idea that a rabbit, say, could acquire language if it were to receive the right stimuli and training. So far, then, we don't have any argument for the claim that the human child begins with something specifically linguistic. Even so, the observation is not bland in its consequences for a theory that seeks to accommodate it.

The general observation depicts the child as moving from an initial general state that covers all potential languages to a specific state that covers just one language. The *descriptive adequacy*, therefore, of a general theory of linguistic competence would appear to involve a delineation of the seemingly infinite variety of languages upon which a child may fixate. On the other hand, if our general theory is to be *explanatorily adequate*, then we need to explain how a child may fixate on any point in this infinity without any such point being favoured prior to the child's exposure to language. In themselves, these constraints leave open the questions of what the child's initial state is, how specific, complex is its final state, and what data the typical child is exposed to such that it can move from one to the other. Imagine, though, that there were only one human language HL, as if we all spoke English. In this scenario, our innate equipment could only support HL – there is nothing else for it to support. Our first two questions are trivially solved: the required generality of the initial state would reduce to the specificity of the final state, the one universal language. In this light, the third question of how a child might move from its initial universal state to a specific end state would also be effectively solved: the only language our mind/brains are able to represent is the only one there is to acquire; thus, the child would need no data to select or decide between languages. The pursuit of descriptive adequacy, however, surely leads us to the thought that the languages we speak are very different. There might be universal features shared by all languages, but they are not apparent in the seemingly infinite variety of data to which children are exposed. But what else does the child have other than the data? It *seems*, in other words, that we are infinitely far from the explanatory ideal situation, i.e., the more languages there are, the more inclusive must be our initial capacity to represent language. Data wise, this means that the child appears to require finer and finer data so that the particular target language may be distinguished from the indefinite other ones it has the capacity to entertain. But here's the rub! The linguist has as much data on the grammar of English, say, as he could wish for; he also has the capacity to reflect upon it, theoretically or otherwise, and the advantage of comparing it with data from other languages, but he still cannot figure out the grammar of English – that's what, inter alia, we have linguistics for! If, then, we content ourselves with the bland remark about nativism, we are led to think of the child who successfully acquires English has having enough data to figure out what self-reflective linguistic inquiry has been banging its head against for the past couple of millennia. Something is wrong.

If our brains are not wired, as it were, to acquire HL, if, instead, we have to find our way to one of an indefinite number of possible languages

from the data we are given as children, then how do we all achieve this feat without enough data? If we are led to conclude that what we do as naturally as walking or eating is impossible, then we know our reasoning is awry. The obvious flaw in our reasoning so far is our implicit assumption about the start point of the child's journey. The bland remark that the child begins with innate equipment is true enough, but we seem to require something decidedly less bland. For example, if what is innate must be complemented with rich data about each construction we are competent with, then, contrary to fact, the child would not acquire the language, for no child is exposed to each construction with which it attains competence. Likewise, if our innate equipment were merely to construct a language out of a given sample based on general principles of pattern or frequency, then, again contrary to fact, it would not arrive at the target grammar, for well-formedness appears not to be based upon statistical or substitutional relations. What the child's innate equipment is required to do, it seems, is actively constrain its 'choices' as to what is part of the language to be attained. But no child is wired to target any particular language: the child can make the right 'choices' about any language with equal ease. This suggests that children must begin with 'knowledge' specific to language, i.e., the data to which the child is exposed is 'understood' in terms of prior linguistic concepts as opposed to general concepts of pattern or frequency, say. If this is so, then we can see how a child may acquire a language even though the data itself is too poor to determine the language: the child needs no evidence for much of the knowledge it brings to the learning situation. In crude terms, children always make the right 'hypotheses' as a function of their genetic endowment. Thus, since the child can fixate on any language in the face of a poverty of stimulus about each language and all languages are equally acquirable, children all begin with the same universal linguistic knowledge. This is the poverty of stimulus argument.

The argument, please note, does *not* tell us: (i) what information is innate; (ii) how the innate information is represented in the mind/brain; or (iii) whether the information is available to a general learning mechanism or specific to a dedicated one. These issues are to be decided by the normal scientific route of the testing and comparison of hypotheses. The first issue depends on what linguistics tells us about the sameness and difference of languages. The second depends, ultimately, on our investigations of the brain and the kinds of cognitive architecture it may realise. The third is especially slippery: it is at least a priori coherent to think of a general intelligence exploiting innate domain specific information. One way of investigating this question is to test the counterfactuals – discover the extent to which putative domain specific competencies may be selectively spared

or impaired, and, more generally, be ontogenetically and synchronously autonomous. If linguistic competence can be attained and maintained in relative isolation, then a general intelligence would be an explanatory dangler with respect to the competence.

It also bears emphasis that the argument does not rely on claims as to the relative availability of positive and negative data. Positive data tells the child that some construction is acceptable; negative data tells it that some construction is unacceptable. As we shall see, there is much discussion of this difference, for it has been claimed that negative evidence is typically unavailable and not used by the child even where it is available. This contention is important, for if there is no negative data, then the child will not be corrected if it initially fixates on a grammar that includes the 'correct' grammar as a proper subset, i.e., there would be no evidence to tell the child to contract its language. Thus, it is claimed that children are innately constrained to initially 'chose' the smallest possible language compatible with their positive data. Much of the debate around the POSA, therefore, focuses on negative evidence; as we shall see, Cowie is no different in this regard. The POSA proper, however, does not differentiate between data. As I hope to make clear, even if there is plenty of negative data (which I do *not* think there is), the POSA is not refuted.

Notwithstanding the relative neutrality of the POSA, it does suggest something surprising: the fact that the child can acquire any language without seemingly enough data to do so, indicates, counter-intuitively, that languages are not so different. That is, the innate 'hypotheses' the children employ must be universal, rather than language particular. This is easy to see. Imagine that each language is radically distinct, an effect of a myriad of contingent historical and social factors. This seems to be what the pursuit of descriptive adequacy tells us. Now, if this were the case, then the child's data would still be poor. But how would innate knowledge help here? Since, ex hypothesi, each language is as distinct as can be, there is no generality which might be encoded in the child's brain. That is, the child would effectively have to have separate innate specific knowledge about each of the indefinite number of languages it might acquire. Yet this is just to fall foul of the POSA again: how does the child know that the language it is exposed to is a sample of grammar G as opposed to any of the other grammars? Here we reach a curious pass. It *might* be that humans are after all in a near *ideal explanatory situation*, such being why they need so very few data relative to what is acquired. This thought is at the heart of the *Principles and Parameters* (P&P) approach which came into focus in the late 1970s. The thought, of course, is not that we are really approximate HL speakers; rather, from the optic of the child's brain, there is simply

language. So, there is massive variety in the languages we speak, but this is analogous to the range of dialects that exist within the one language, such as English. The specific conjecture is that we all begin with *universal grammar* (UG), the one language, as it were. UG is innate and is unformed in the sense that it encodes certain options or *parameters* which are set by exposure to certain data. To acquire a language is simply for the values of UG's parameters to be set in one of a finite number of permutations (given the acquisition of a lexicon). The choice of permutations ramify to produce a seeming infinite variety. Let me say something very brief about this picture as it is currently understood, then we can move onto Cowie.

Chomsky (e.g., 1995) understands UG to be the initial state of the *language faculty* (an abstractly specified system of the brain). Assume that the faculty is situated within an ensemble of cognitive systems, primarily, a sound-articulation system and a conceptual-intention system. The faculty comprises a lexicon, which we may take to be a list of exceptions, clusters of features idiosyncratic to particular words; these include, minimally, those features which enter into the interpretation of the faculty's outputs at its interfaces with the other systems of the mind: categorical features ($\pm$V, $\pm$N), phonological matrices, agreement features (for N), and semantic features. Think of the faculty as realising a computational procedure C_{HL} that maps a pair of output representations (structural descriptions) onto a selection from the lexicon. The representations are respectively a phonological form (PF) and a logical form (LF), a pair $<\pi, \lambda>$; say that the computation *crashes* if the two representations are not *compatible*, where compatibility is determined by the conditions imposed by the faculty's interfaces with other systems; i.e., the output conditions are such that a PF representation can be *read* by the sound-articulation system and a LF representation can be read by the conceptual-intention system. Such conditions, in other words, guarantee that the representations are legible to other systems (*not* to the preclusion of gibberish).[3] Simply put, to acquire a language is to acquire a particular systematic mapping between sound and meaning. How do we fixate on such a pairing? Think of the language faculty as being in a genetically determined *initial* state S0 prior to experience; what experience *triggers* is the setting of values along certain parameters (perhaps just on-off) that determine the output conditions for $<\pi, \lambda>$ and, of course, experience provides the assignment of features in the lexicon, although not the features themselves. Different experiences set the parameters to different values; this finite variation ramifies to produce languages of seemingly infinite variety. Once all parameters are set, the faculty attains a *steady state*, Ss, this we call an I-language: that generative system which explains an individual's competence with his idiolect.

To repeat, this kind of story of UG is not at all implied by the above general reasoning about acquisition in the face of poverty of stimulus. It is, rather, a somewhat speculative hypothesis based upon a myriad of considerations, both empirical and theoretical. Indeed, the form of the POSA is quite general and based on what Chomsky (1986, 1991) has called *Plato's problem*. The problem occurs wherever a competence is exhibited which we have apparently too little data to acquire. The slave boy in the *Meno* was not taught geometry, but he still acquired knowledge of Pythagoras' theorem. Just so, we are not taught the many principles and constraints which essentially enter into an explanatorily adequate theory of our linguistic competence, but we acquire the competence still. Any such consideration only militates for our antecedent possession of some specific knowledge to take up the slack between available data on, and mature competence with, mathematics, language, etc. The hypothesis that the slack is taken up by *such and such* knowledge is a distinct conjecture which is as hostage to empirical fortune as any other hypothesis. Thus, in the language case, the POSA is *not* employed in *direct* defence of UG (under some proprietary specification); rather, UG is supported to the extent that it is the best theory of the knowledge which the POSA tells us exists (see Matthews, 2001).

Let us turn to a familiar concrete linguistic example, much as Plato turned to Pythagoras' theorem, and examine Cowie's sceptical assessment of it. It should be stressed that any particular example, *qua* datum, is at best suggestive. Cowie appears to think that POS considerations reduce to a few examples; the shallowness of this conception will emerge as go along. For the moment, let us play by Cowie's rules.

Cowie's chief example is the familiar case of polar interrogatives (questions which may be meaningfully answered with 'Yes' or 'No'). Let us assume that the typical (English) child has only the data in pairs such as (1)+(2) from which to determine the correct interrogative form:

(1)a.   That man is happy

   b.   Is that man happy?

(2)a.   That man can sing

   b.   Can that man sing?

What must a child know such that it can correctly go from this kind of data to the correct interrogative form in general? Chomsky (1975) asked this question as a challenge to Putnam (1971), who had contended that

the child need only have at its disposal general principles (not domain specific linguistic ones). Cowie (1999, 178), after Chomsky (1975, 30–31) considers the rule SI.

(SI)   Go along a declarative until you come to the first 'is' (or, 'can', etc.) and move it to the front of the sentence.

SI is *structure independent* in that it appeals merely to the morphology and linear order of the declarative. The important point here is that an empiricist may happily appeal to SI as the rule upon which the child fixates, for it involves no linguistic concepts and so is one at which a child may arrive without the benefit of specific linguistic knowledge. Now the child would proceed correctly with SI so long as she continued to meet such monoclausal constructions as (1)+(2). But the rule does not generalise. Consider:

(3)   That man who is blonde is happy

Application of SI would produce the nonsensical

(4)   * Is that man who blonde is happy?

Would some other, more complex rule suffice? Well note that we have already gifted the empiricist an identification of verbs that carry agreement, modal and aspect features, but for our purposes an empiricist is someone who rejects any specifically linguistic knowledge prior to experience; so, the empiricist cannot help himself to the notion of, as it might be, inflectional head. The empiricist, it seems, needs some antecedent rule to make this identification. I shall return to this point shortly. A rule which is sufficient to cater for (3) is

(SD)   Move the auxiliary verb which occurs immediately after the subject NP to the front of the sentence.

This rule generalises to all cases so far considered (as we shall see, SD is quite inadequate as a generalisation. The present concern is just to consider the kind of features of a rule that has at least some hope). The key feature of SD is its *structure dependence*: an adequate rule must advert to the auxiliary and the subject NP; that is, it must be sensitive to the *hierarchical* phrase structure of the sentence, e.g.:

(5)   [NP That man [CP who is blonde]] is happy

(The form here is very simplified; the auxiliary is usually treated as an inflectional head of a projection IP. The rule at issue must thus advert to *head-to-head* movement under which the auxiliary moves from I head to C head of the 'highest' CP. This kind of movement may be subsumed under a more general constraint pertaining to *shortest movement* and *agreement checking* (Chomsky, 1995). Such subtleties will, in the main, be forsaken for purposes of exposition.) Now if we assume that the child only has the data recorded in (1)+(2) to go on, then she appears to make a massive leap to the auxiliary inversion rule SD. After all, SI predicts the pattern in (1)+(2), so if the child had no antecedent information on phrasal structure, why on earth should she opt for the complex and particular rule SD? But she does insofar as SD, but not SI, is descriptively adequate (up to a lose measure). The hypothesis that the child has specific knowledge of language prior to experience makes sense of the child's 'choice' insofar as it claims that the child's mind is wired to make definite choices as to the distribution of agreement features and to select a particular option for the realisation of those features in the output representations of the language faculty. In other words, SD is simply an epiphenomenon of the triggering of the language faculties of children exposed to 'English'. The obvious response an empiricist might make here is just to say that the child *does* have enough data either to target SD directly or to falsify SI and move onto SD; our restriction of the data to the pairs in (1) and (2) is mere theory laden stipulation. Such is Cowie's tactic.

Cowie's (1999, Chap. 8) bold claim is that empiricism, whether of the Putnamian or enlightened stripe, is untouched by poverty of stimulus considerations. She supports this contention with two substantial points. First, the data available to the child (hereafter, primary linguistic data – pld) is much richer than is usually thought. In particular, if the pld contains negative evidence, then the child could 'overshoot' but still arrive at the target language through correction. Second, the empiricist need not be saddled with obviously inadequate postulations such as SI; the empiricist (of an *enlightened* caste, at least) can coherently appeal to structure dependent rules. I shall tackle each of these points in turn and then offer a diagnosis of, and corrective to, Cowie's arguments.

Cowie (1999, 179) thinks that a key premise of any poverty of stimulus argument is of the form that, for any given 'rule', such as auxiliary inversion, the pld is too poor to allow a generalisation to the rule without antecedent knowledge of the structure. This is then taken to "license the Chomskyan position" and so refute the empiricist. Cowie questions whether the pld is so poor. She produces a roll-call of quotations which suggest that Chomsky simply assumes that "no evidence exists that would

enable a three-year old to unlearn an incorrect rule" (1999, 186). But, appealing to Sampson's (1989) and Pullum's (1996) data, Cowie thinks that the pld might be sufficiently rich to allow children to move from SI to SD. Pullum, for example, estimates that, within the typical pld, confirmatory data for the correct rule of auxiliary inversion make up 1% of interrogatives and 10% of polar interrogatives. Cowie's (1999, 187) judgement is that it is "just bizarre to suppose that children encounter *no* sentences of the kind that would disconfirm [SI]". She concedes, of course, that Sampson's and Pullum's data is only suggestive; even so, given the lack of good data on the pld, her thought is that it is really too soon to make the kind of claims Chomsky makes; that is, we just do not *know* how poor the pld is. This being so, perhaps the child, as the empiricist claims, is free to make false hypotheses, such as SI, and have them falsified by the pld. So long as this remains a possibility, there is no sound poverty of stimulus argument, or so the thought goes.

I agree with Cowie that it would be surprising if *no* child were exposed to data (positive or negative) that falsified SI. At the very least, I accept that we do not know enough about the pld to infer directly that the child is innately armed with the resources to frame SD. I also agree that many 'Chomskyans', as her quotations attest, make ill judged claims about the pld. Cowie is surely right: we just don't know enough about the pld for it alone to rule in or out any theory of the child's mind. However, the ill judged claims are such, not so much because they are clearly wrong, or even likely to be wrong, but because they are misleading. It is, I think, abundantly clear that the logic of the nativist POSA does *not* depend on every child lacking this or that datum. It is, rather ironically, the empiricist who makes unsupported – "bizarre" – assumptions about the pld. Let me explain.

Chomsky's real point has never been that the pld just doesn't contain this or that construction. It is true that Chomsky has made assumptions to this effect (see Cowie, 1999, 184), but such claims are always in the context of a challenge to the empiricist. The importance of this background is that Chomsky may be properly understood as arguing that the pld is not as it *must* be if empiricism is to be true. Consider again the general reasoning laid out at the top of this section. The argument does not rely on particular hypotheses as to the data children encounter; rather, it goes through on the *general* thought that were a child learning a language with no specific antecedent knowledge, the data required would be absurdly rich given that the child, cognitively speaking, would need data on each choice which differentiates its target language from every other language it might acquire. Indeed, further, the child would need data to rule out

choices which lead to no language. Such was the point of the thought that linguistics itself does not seem to have enough data to fixate on the grammar of English, still less the universal principles realised in UG, even though linguistics has potential data on every language. The real point of the POSA, then, might be construed as a generalisation of the of SD example; i.e., it is unreasonable to assume that, for a child to fixate on rule R, it needs exposure to all the distinct types of construction to which R applies, i.e., all those construction types which would refute potential prior hypotheses of 'false' rules. This reflection is striking where just one child is being considered; for if current linguistic theory is anywhere near correct, and Cowie does not doubt that it is, the properties of all constructions are accounted for by quite simple abstract principles. It thus seems that if the child is to fixate on the right principles, then it must be exposed practically to every possible construction (some indication of this will be given below). When, however, we consider populations of speakers, the pld must not only be super rich but also uniform. Well, we might not know that much about typical pld, but it is patently not uniform in the sense of every child's pld including every potentially falsifying construction up to the correct principle. So, unless children all miraculously jump to the same correct hypothesis, we would rationally predict massive variation in their mature rules simply because their hypotheses would be based on different data. And even if we could make this convergence seem less miraculous, we would still want to know why there is any convergence at all. There is, of course, variation between speakers (the very notion of an *I*-language is intended to capture that), but it is highly restricted (according to minimalism, all variation is morphologically driven due to lexical differences – the 'syntax' is univocal). For example, as far as we know, the essence of SD is a descriptively adequate generalisation of polar interrogatives (which does not mean that a more encompassing generalisation is not to be sought) over every 'English' speaker, but SI is not. Yet to arrive at SD requires data richer than are required to arrive at SI. Here we see that the focus on particular constructions is expository useful but highly misleading. The POSA runs on what data we can expect children in general to possess, not specific children-construction pairs. We will *never* be in a position to be dogmatic about that, and no-one has ever thought that we might be.

In sum, POSA does not assume that the pld is just *so* and *thus*; rather, it rides on the back of the thought that the pld is poor relative to what it must be like if children were learning a language from the degraded and variable data. This, I take it, is Plato's point in the *Meno* (although I'm no scholar): there is nothing in particular being withheld from the slave boy, but he arrives at an understanding of Pythagoras' theorem on the basis of

data that would not be sufficient were he relying on just that data. Hence, we conclude (*non*-demonstratively) that he has prior knowledge about the domain. Just so, the linguistic nativist's POSA does not say that children just do not encounter polar interrogatives with relatives attached to their subject NPs; rather, it says that, given the specific complexity of what the child does acquire, there is good reason to think that the child could not arrive at his target language were he to rely on data alone. After all, we have all the data we could wish for but still cannot arrive at (a theory of) what the child knows. This leads to the thought that the child has innate domain specific linguistic knowledge to take up the slack. Pretty plainly, I should say, as regards the pld, the nativist has the *weaker* hypothesis, for it obviously does not require the pld to be necessarily poor, only not super rich. It allows, therefore, for great variation and all manner of degradation and ambiguity in the pld. The empiricist hypothesis, on the other hand, appears far too *strong*, for it apparently requires the pld to be richer than the linguists' data! For instance, even assuming that all children are exposed to polar interrogatives with relatives attached, the empiricist child, if starting with SI, would need *other* data to tell it that the constructions were in fact grammatical. After all, even if the child were to find itself reading the *Wall Street Journal* or Oscar Wilde, if it were initially working with SI, why would it not dismiss (3) as anomalous? It would seem that it is the empiricist, not the nativist, who is wedded to "bizarre" assumptions about the pld.

On reflection, this point becomes obvious when set in the wider context of the UG model set out above. On the nativist view, the pld need only contain that minimal information that is sufficient to trigger the parameters on the realisation of the principles of UG, any other data available to the child makes no difference. Lightfoot (1993), for example, has suggested that all parameters are set by 0-embedded clauses (e.g., (1) and (2)).[4] Even if this particular conjecture is incorrect, the general idea seems to be right. The data which trigger a parameter *must* be plentiful in the pld, because all children, barring pathology, do have their (marked) parameters set. That is, the less plentiful the triggering data are, the more likely it is that some children would not arrive at their target grammar. As it is, *all* children do fixate on their core target grammars and tend to make the same 'errors' along the way. Thus, the general nativist position does not depend upon tendentious assumptions about the poverty of the pld, but only upon the thought that the pld can be very poor (no negative evidence, say), for the realisation of a mature grammar (an I-language) from UG depends only on the most plentiful constructions, those to which all children are exposed with no negative feedback. There is no such give in the empiricist position,

which is obliged not only to show that, barring miraculous coincidence, the pld contains *all* constructions (whether positively or negatively framed) covered by the right grammatical generalisation, but also that children make use of them, i.e., the presence of such constructions is necessary for the fixation on the 'correct' interpretation. The burden of proof about the pld is thus squarely with the empiricist. Well, is the pld rich, and does the child require it to be such?

Cowie's claim is that the pld might be so rich as to permit error correction, a falsification of SI. Cowie's Chapter 9 is largely given over to questioning the 'assumption' that there is no negative evidence in the pld, such being why the pld is thought to be so penurious. The key research here is that by Brown and Hanlon (1970), which found, in essence, that mothers do not explicitly correct their children's grammatical errors. The conclusion was drawn that children do not make use of negative feedback. This conclusion has since been challenged by Bohanon and Stanowicz (1988), Demetras et al. (1986), Hirsh-Pasek et al. (1984), and Moerk (1984). The common theme of this research is that while mothers do *not* tend explicitly to correct their children, they do offer a variety of cues to the child, such as repetitions, questions, and recastings. On the basis of this challenge, Cowie (1999, 230) avers that Brown and Hanlon (1970) is an "outdated shibboleth". This counter-conclusion is really quite inappropriate.

The data Cowie appeals to has *not* been taken to intimate a rich source of negative evidence (e.g., Penner, 1987; Bohanon et al., 1990; Gordon, 1990; Marcus (1993); Morgan and Travis (1989); Grimshaw and Pinker (1989)). The kind of negative evidence putatively exploited by children is very weak, only appears in mothers with young children, and was not even found in all the dyads studied by Bohanon and Stanowicz (1988). Crucially, the relatively rich mother-child interaction observed is typical of the Western middle-class, but it is far from universal (especially see Heath, 1982). Cowie's (1999, 232–234) response to this challenge misses the point. She protests that even if the negative evidence is not universally available, it does not stop the children who have it from exploiting it; different children may exploit *different* sources of evidence. Well maybe, but that is certainly not shown by the presence of very weak negative evidence within a particular culture. The *fact* that children acquire normal competence without such evidence *shows* that the children who do have it do not need it. This is corroborated by there being no correlation between negative evidence supplied by an attentive mother and faster acquisition of mature competence (Newport et al., 1977). So, not only do children not require the negative evidence, even when they have it, they don't use it (this observation is also supported by a wealth of anecdotal data on the

sheer recalcitrance of children's errors). As we shall see later, Cowie tends to overplay the significance of at the very best suggestive data because she insists on attributing to the 'Chomskyan' dogmatic and even a priori claims which in fact are not held.

All the data we have indicate that children's errors (morphological, semantic, syntactic) are quite rare, certainly rarer than they would be were the child seeking to falsify or test initial hypotheses. Moreover, the errors made are neither random nor occur equally for all constructions. For example, familiarly, children (as well as adults, of course) make regularisation errors with the past tense affix -ed (although quite rarely in fact), but never has it been attested that they regularise auxiliary verbs.[5] Likewise, young children make doubling errors, where a phrase (*wh* or auxiliary) occurs at an extraction site as well as it its landing position; this tends to occur only where the two positions are separated by a long clause with a transitive verb and its complement; that the doubling error does not appear otherwise, suggests that children are not suffering under a 'false rule', but are, rather, making performance errors (Nakayama, 1987). It is, however, very difficult to talk sensibly about children's errors in the absence of an acquisition model, for, whether rare or legion, the pattern of errors remains unexplained. This simple point is unfortunately often neglected. A theory of language acquisition must explain what we get 'right' just as much as what we get 'wrong'. Just so, as the specific complexity of our competence leads to a theory of UG, so the specific systematicity of our errors leads to the thought that we are not, in general, falsifying hypotheses. Cowie is apparently oblivious to any such consideration, for she seems to think that, egregiously, the mere existence of errors militates for empiricism, or, rather, some as yet unspecified learning regime based on general principles.

Cowie (1999, 185) employs the Sampson/Pullum data to combat Crain's (1991) finding that children do not make errors consistent with SI, i.e., they do not produce nonsense like (4). Now, if one is attempting to support at least the feasibility of a generate and test model that is only generally constrained, then, minimally, one needs to show that: (i) the errors children do make are best explained as false hypotheses about the target grammar; and (ii) there is the data in the pld to falsify them; i.e., falsifying data is necessary for the language learner. The Sampson/Pullum data is at best suggestive about (ii), it says nothing about (i). The importance of the first condition is that if children's errors are systematic in the way indicated, then it is inadequate to view them as consequences of generally constrained hypothesis formation; in the limit case, if children don't even make the errors at issue, then there is simply no role for falsifying

data to play. Such is why (i) and (ii) are minimal requirements; without their satisfaction a debate about empiricism is empty. Otherwise put, their satisfaction would only show that empiricism is not *dis*confirmed, *ceteris paribus*. Again, because Cowie's eye is on a level of confirmation that appears to be quite unscientific, she fails to connect with the real issue of how errors may be explained.

Cowie (1999, 201–202) does no more than cite some of the peer commentary on Crain's (1991) BBS target article to the effect that some of the children did make "the supposedly impossible errors". But Crain (1991) already concedes the presence of errors; neither Crain nor, to my knowledge, anyone else consider errors to be "impossible". The fact is, though, that Crain's data and those of Nakayama attest to not a single error which is suggestive of the employment of a structure independent rule; certainly none of the children produced (4) and its like. This is quite striking given that a range of other errors were committed. The crucial issue is how errors are explained, and there are many ways of classifying and explaining errors that are perfectly consistent with the nativist stance. Crain and Thornton (1998), for example, uphold the *continuity hypothesis*, which says that a child's errors should be consistent with some parametric value of UG, i.e., the errors are only such relative to the target language, not UG (cf., Pinker, 1984). On the basis of this hypothesis, errors may be explained by features of the test sentences (e.g., failure of the experimental situation to satisfy presuppositions of the test sentences, thus creating pragmatic difficulties for the children) as well as more general extra-linguistic factors, not only the non-linguistic cognitive demands an experiment might make on a child, but also simply poor controls, as well as unavoidable 'noise'. Crain and Thornton (1998), in the light of such explanatory options, hypothesise a *modular matching* acquisition model that claims that the child matches the adult in processing mechanisms and UG principles; the model predicts that once all such factors listed above are properly controlled for, children's performance should be effectively flawless (up to 90% mean) relative to the constraints of UG (e.g., structure dependence). This is not "dogma", but a real research strategy with much supporting data. It bears emphasis, however, that one need not adopt such a hypothesis to remain a nativist of the Chomsky stripe; many acquisition models are consistent with the existence of UG. It takes not a little discipline to ignore all of this.

On the most charitable reading, then, Cowie shows that there *might* be the data in the pld to falsify SI, but this we knew anyhow. The empiricist needs the separate and much stronger thesis that the pld is sufficiently uniform in its containing the crucial data *and* that the child does in fact make use of it. Cowie does not even attempt to support this joint thesis.

Much of the above should be academic, for no-one serious thinks that consideration of the pld, or assumptions thereof, pertaining to particular constructions will be decisive one way or the other. When theorists do speculate on the pld, they should be read as speaking of the pld in the round. That is, *the pld is as variable as it is poor*, but children still achieve the same level of competence within bounds that would suggest a uniform pld were the children hypothesis testing. I hope to make this point vivid with respect to the auxiliary inversion rule when we turn to diagnosis, but let us first consider Cowie's claim that the enlightened empiricist may appeal to grammatical structure.

Given my characterisation of the standard reading of the auxiliary inversion rule, Cowie's appeal to a rich pld might seem beside the point. Recall that Cowie is interested in showing that the POSA does not refute an empiricism that renounces domain specific information; the child might get by with *general* purpose strategies. Yet the correct rule, the one on which the child must fixate, essentially involves syntactic concepts, it is structure dependent. Whether or not, then, the pld contains data which falsifies SI, the child will not arrive at SD if it lacks the requisite syntactic concepts with which to formulate the correct rule as an hypothesis (here I am assuming an empiricist hypothesis testing model). Cowie (1999, 189) considers this objection as if its purpose were to buttress the creaking assumptions about the pld. This is not so. Cowie's polemical agenda aside, the objection is clearly of a piece with the general conjecture that the knowledge that constitutes a mature competence is not such as to be derived or induced from the pld. But is this thought correct?

As with the above considerations about the pld, Cowie's tack is to present some data which are taken to indicate that innate linguistic knowledge is not required by the child; (enlightened) empiricism is not refuted. Again, because Cowie thinks that Chomsky and others peddle the POSA as a knock-down argument against empiricism, she presents the data as if it were a refutation of an a priori claim. Data is data is data. The point, as with any science, is to weigh the data in the round and see which explanatory theory is best corroborated. Philosophically, this might seem slightly dull, but such is the only serious way to proceed. As it happens, however, the data Cowie (1999, 190–193) discusses is beside the point.

The nativist hypothesis before us is that the child requires specific knowledge about syntactic structure to fixate on structure dependent rules; no mechanism, dedicated to language or not, constructs or extracts the structure from the pld. Now the only data relevant to this claim will be that which, not only bears on syntax, as opposed, say, to morphology or

phonology, but is also developmentally realistic. Cowie's discussion fails on both counts.

Cowie's first claim concerns 'Motherese'; her hypothesis is that the pauses, exaggerated intonation, and distinctive prosody of a mother's speech to her child will give it a platform from which to extract the syntactic categories from the phonetic stream. As Cowie readily admits, the 'Motherese' hypothesis is out of favour. The reason for this, Cowie (1999, 190) suggests, is that it has been construed as the claim that the child constructs syntactic categories from prosodic properties. But such is not Cowie's claim; for her, Motherese provides an initial framework from which the child may proceed to abstract statistically syntactic categories. I beg to differ. The unpopularity of the Motherese hypothesis has two principal sources, unmentioned by Cowie. First, Motherese is not a universal phenomenon: some cultures and communities either lack Motherese all together – parents speak to their children with no peculiar prosody – or parents actually tend not to talk to their children much at all; even so, the children acquire their respective languages perfectly well (Heath, 1982; Schiefflin and Eisenberg, 1981), including, of course, deaf children (e.g., Feldman et al., 1978; Newport and Meier, 1985). Second, differential exposure to Motherese is not correlated with differential rates of language acquisition (Newport et al., 1977). Whatever Motherese is *for*, it does not appear to have a decisive role in language acquisition. What, however, of the specific claim that the child may use prosody to extract syntactic structure?

Glietman and Warner (1982), to which Cowie (1999, 190–191) appeals, offers no support to Cowie, for the study just does not concern syntactic structure, but rather word segmentation. If sound, the data does indeed indicate that prosody potentially provides the child with "important information about word boundaries" (Cowie, 1999, 190). But so what? The issue is about phrasal boundaries of the kind that are essentially involved in the auxiliary inversion rule. No linguistic nativist need be committed to saying that word individuation is achieved independent of phonetic and contextual cues. So, where is the data that the child can build on the statistical segmentation of words to arrive at phrasal boundaries? Cowie does not provide any. She speculates, following Marastsos (1982), that the child, once having segmented a class of words, such as *Daddy*, *doggy*, etc., will be able to recognise them as occurring in the same position in some simple matrices which marks them as subject NPs. Cowie also lays great store by the study of Saffran et al. (1996) which found that children under twelve months can distinguish sound sequences spanning word boundaries from those that form words. It thus appears that children have quite developed

mechanisms for extracting statistical regularities, mechanisms that appear to be general and do not involve syntactic categories; so, the mechanisms are open to the empiricist. Again, the implication is that with such sensitive machinery the child should be able to extract syntactic structure. Well, maybe, but the data do not lead us to that conclusion.

Prosody, especially that of Motherese, *might* reflect word boundaries, but it is far from clear if phrasal boundaries are reflected (Pinker, 1987). In effect, then, what the child must be able to do, if she is to progress from words to phrases, is recognise that *Daddy*, as it might be, is the *head* of a subject NP, but this is something that looks not to be either phonetically or morphologically marked. I am not suggesting that the child does not analyse (parse) its input stream; my point, rather, is that to do this the child requires some structural constraints (phrase bracketings) specific to language and there is no data to suggest that this is encoded in the input. Which is not say, of course, that prosody might not contain many cues for the correct bracketing. Much of Cowie's thinking, I believe, founders on this issue. She moves from the pld containing cues to phrasal boundaries (a thought which offers no significant support to empiricism) to the pld containing the phrasal information in a statistically recoverable form (a thought which amounts to empiricism). (See Gleitman, 1990, for a subtle corrective to this kind of leap.)

Cowie understands the data she has marshalled to indicate that there are potentially enough cues in the child's environment to support the abstraction. She (1999, 192) claims, for example, on the basis of the Saffran, et al. study, that the child's statistical wherewithal can perhaps exploit semantic cues (effectively, θ-categories of agent, patient, etc.) to infer syntactic categories. Although Cowie does not make it clear, this idea is essentially that of "semantic bootstrapping" as proposed by Grimshaw (1981) and Pinker (1984, 1987). The idea is certainly worth pursuing, but it does not militate for Cowie's claims. First, the θ-roles are understood to be innate. (The child has to hypothesise, as it might be, 'All *objects* are named by count nouns'. Where does *object* come from? See Fodor (1998, Chap. 3).) Second, the bootstrapping mechanism need not be understood as a property of UG; it may, rather, be construed as a separate mechanism that maps semantic properties onto the syntax proper. Thus, there is nothing in the idea that challenges any core claim of the Chomsky position. Third, bootstrapping offers no reason to favour a statistical model of learning rather than a rule-constraint based one. In line with enlightened empiricism, however, Cowie (1999, 226, fn. 20) suggests that we may think of the child as discovering constraints as she goes along rather than the constraints being selected from an innate store; she neglects to explain, though, from where

the content of the constraints derives. If they are induced from the pld, then we are back to square one: we want to know how the (empiricist) child can discern θ-role assignment conditions in the pld just as we want to know how the child can discern phrasal structure. Equally, if the constraints are not induced from the pld, then no evident progress has been made. It is just this kind of explanatory lacuna that makes nativism so attractive.

Recall that the purpose behind Cowie's presentation of the above data is to cast doubt on the anti-empiricist claim that syntactic categories are not recoverable from the pld. What the data in fact shows is that children have quite sophisticated statistical abilities. Now this would have real significance if the linguistic nativist were peddling an argument with a premise that denied such abilities. But, notwithstanding the exigencies of Cowie's polemic, there just is no such argument. Indeed, the abilities, for all we can tell, *are* specific to language, as Cowie herself concedes.

Cowie is not arguing for empiricism, she is simply suggesting that it is not refuted. Well, sure; certainly no-one should reject research that has not been carried out. But such a modest contention sits uncomfortably with Cowie's (1999, 193) hyperbolic conclusion:

> there is no reason to believe that a[n] empiricist would necessarily get hung up on the false rule [SI] ... [T]here also is good evidence that [the empiricist is] perfectly able to acquire the 'abstract' syntactic concepts that they need to form hypotheses through statistical analysis of the speech they hear around them. [The poverty of stimulus argument fails to] demonstrate the falsity of [empiricism].

The last sentence is certainly true, but, to say again, there just is not a demonstration on the cards, one way or the other; there are empirical and theoretical considerations as in any other science. That Cowie thinks something different is at stake perhaps explains the confusion in the first two sentences. Of course, the empiricist is not *necessarily* saddled with SI. But some work needs to be done to *show* how a general mechanism may arrive at SD; it is not good enough to talk vaguely of a mechanism that has a "preference for rules stated in terms of unobservables over those stated in terms of observables" (Cowie, 1999, 189). It is not as if any old unobservables will do; the constraint is quite specific. We want to know specifically how the child can have a "preference" for 'rules' involving subject NP and matrix auxiliary verb (as indicated, and as we shall see below, the constraint actually operative is much more abstract). Cowie's hand waving would be appropriate, though hardly satisfying, if the point at issue were an a priori one, but the question is straightforwardly empirical. Thus we turn to the putative "good evidence". There *is* evidence that the child is able statistically to recover some information from phonetic streams, but there is *no* evidence that the child can statistically induce syntactic categories.

None of the authors cited understand themselves to have *good* evidence for *that*.

A prevailing theme of the above discussion is that Cowie misreads the theorists of the generative tradition as seeking a *demonstration* that any form of domain general learning mechanism is inadequate to fixate on rules that are in essence syntactic. Cowie, to be fair, does have Pullum's (1996) reconstruction of the 'Chomskyan argument' in mind. Pullum presents the argument so as to refute it, but Cowie (1999, 196, n. 21) finds it an "irresistible target", for it is "so much more clearly and forcefully stated than [the] nativists' own versions". The nativists' versions are not "clearly and forcefully stated", I have suggested, because no-one serious is interested in knock-down arguments; there are certain empirical and theoretical constraints and a substantive proposal to satisfy them. Our being told that the proposal is not a priori true is not news, especially to Chomsky. In itself, Cowie's myopia on this point is perhaps of no great significance, yet it leads her into a fundamental confusion about the methods and goals of linguistic theorising.

Cowie (1999, 197) thinks that the auxiliary inversion rule is treated as a *experimentum crucis* in the literature. Further, Cowie understands the linguist as responding to the inadequacy of the auxiliary rule alone to prove his case by producing "other cases, involving different grammatical rules and principles" (e.g., want+to contraction, principles of binding theory, etc.) that are similarly claimed to be unlearnable. The obvious problem with this proposed nativist methodology is that, like the "many-headed Hydra", rules are produced and then cut-off by the empiricist, to be replaced with different ones, and so on and on. Cowie (1999, 197/201) appears to think that, while this situation is philosophically unsatisfying, it is just part and parcel of a naturalistic approach to language and mind, with the problem for the nativist being that such a piecemeal strategy cannot demonstrate the truth of his position over empiricism.

Not only is this model of the situation inaccurate, but also, it gives the empiricist a much easier ride than she has or deserves.

To begin with a historical point, the "obsession", as Cowie has it, with particular rules is characteristic of *Syntactic Structures* (Chomsky, 1957) and its development into the *standard theory* (e.g., Chomsky, 1965). The *Principles and Parameters* (P&P) approach, developed in the mid-1970s, and its *minimalist* progeny are precisely marked by their rejection of rules. Indeed, the very ad hoc nature of multiplying increasingly complex rules for each new construction identified in each observed language made the postulated grammars quite unrealistic from an explanatory perspective; for the less constrained UG is, the greater the likelihood of the child over-

shooting, and thus the greater the need for apparently unavailable negative evidence. Moreover, no sense could be made of the question, 'Why these rules rather than those?'; the character of the grammars postulated appeared inexplicable. Rules, as I put it above, are epiphenomena: they are neither formulated, nor represented, nor tested by the learner; nor are they theoretical postulates.[6] We can talk about rules, but only for taxonomic convenience.[7] It is thus simply false that Chomsky or others think of the auxiliary inversion rule as crucial; it is a mere taxonomic effect, whose interpretation and explanation has changed radically over the years. What is characteristic of the generative approach over the past 25 years or so is a search for universal principles that unify disparate phenomena, hence the three principles of binding theory, X-bar theory, ECP, etc., which, in their turn, have succumbed to minimalist pruning, especially the latter two mentioned. Far from one construction-rule pair after another being brought forward, the methodology of post standard theory linguistics is to view the properties of particular constructions as reflecting the interaction of a small number of principles. The details are presently unimportant; the point is that "different rules and principles" are *not* multiplied to challenge the empiricist, quite the reverse. The theory as a *whole* earns its keep insofar as it accounts for the data and meets theoretical constraints (the minimalist program (Chomsky, 1995) pushes the latter condition to shed the P&P model of redundancy). Particular cases and constructions, therefore, are simply data to be accommodated, they are not sought out to refute empiricism. There is no Hydra. If there were, then the empiricist's task would be considerably easier: a single sword swing for a single head, not that the empiricist has managed to decapitate a single 'rule'. Further, the very notion that linguistics is in the business of refuting empiricism is plain silly. As described, the linguist attempts to construct theories that, as in any other science, have universal scope, economy, and predictive success. This endeavour is in itself quite independent of claims of nativism. Indeed, Cowie nowhere disputes a single linguistic hypothesis, certainly not the auxiliary inversion 'rule'. The psychology proper begins when one construes the theories as answers to the question of what speaker-hearers *know*; consequently, the questions are raised as to how we acquire the information and put it to use. Such a construal, of course, places constraints on the theories (*explanatory adequacy*), but these are quite innocent, for there is no a priori bar on empiricist answers to the problems. But the empiricist must now account for the underlying pattern that unifies a host of disparate phenomena, rather than one apparently ad hoc rule after another. I shall exemplify this problem for the empiricist by looking at the auxiliary inversion rule. The thing to note throughout, is that there is no

need to appeal to "other cases, involving different rules and principles", the single case proves the point.

Recall that the auxiliary inversion rule tells us that polar interrogatives are formed from declaratives by the movement of the auxiliary verb over the subject NP to sentence initial position. The point here is not that anyone thinks that children represent SD itself, but that SD describes the kind of concepts and structural relations to which a language acquiring child must, in some sense, be sensitive. Still, roughly, we found that fixation on SD is adequate to deal with monoclausal cases as well as subject NPs with relatives attached. Now does the rule only pertain to these cases of polar interrogatives? Or better, would grasp of the constituent notions of SD allow the child to acquire competence with just the cases so far considered? If so, then Cowie's contention that the innateness of any rule can be disputed seems reasonable. However, the more distinct cases the rule bears on, the much more difficult it is to tell an empiricist story; for, trivially, the less the constructions have in common, the more it is that the only thing they do have in common is the application of the rule, and this has the consequence that *all* children who fixate on the rule would require richer and richer data. (The following discussion is highly simplified. My purpose here is simply to indicate the complexity of the data which needs to be explained, rather than argue for this or that theory of the data. After all, theories can always be disputed, data is more recalcitrant, and it is the data that damns Cowie's characterisation of generative grammar as ad hoc rule stipulation.)

Consider first *wh*-questions (*where*, *which*, *when*, *what*, *why*, etc.), which are not answerable by yes-no. It might seem that given this difference SD would be redundant. After all, if the child is trying to induce syntax from the arrangements of specific words, then why should the child associate *wh*-words with auxiliaries? Consider:

(6)   Which car will Harry steal?

The object of *steal* is *which* car. In the GB approach we would say that *which* car occurs in object position at D-structure and is consequently *raised* to front the sentence (at [SPEC, CP] position) at S-structure. We can, on the other hand, simply note that *wh*-phrases can occur in object position – so-called *echo questions*. Consider being told 'Harry will steal the red car'. Because your hearing is not good, you respond with

(7)   Harry will steal which car?

We do not have to think of (7) as the D-structure form of (6), but the comparison is striking. Moving *which car* effects auxiliary inversion, with *will*

rising above *Harry* to head CP, with the moved *wh*-phrase as its SPEC. As the reader can verify, this is quite general: if the *wh*-object of the main verb occurs at [SPEC, CP], then the auxiliary undergoes head-to-head movement and inverts with the NP position to occupy the position of the head of CP.

The auxiliary inversion rule, therefore, has application outside of polar interrogatives. Of course, there is no real rule; the point is that auxiliary inversion is witnessed in questions generally. It thus seems that the child does not learn one rule for polar interrogatives and another for *wh*-questions; indeed, in itself, the auxiliary inversion rule SD is patently inadequate to capture the distribution of inflectional features. The child's competence appears to arise from a general understanding of the hierarchical relationship between inflection heads (e.g., auxiliaries) and other heads. The data the empiricist must account for is thus much more complicated than at first seems. It seems, for instance, that if the model of polar interrogatives is followed, then the child must learn (6) from (7), but if there is any relation of precedence, it appears to run in the opposite direction, from (6) to (7), in that (7) is standardly used as a response. Moreover, *wh*-fronting from echo questions is not general. Compare:

(8)a.   Harry saw Bill with who?

   b.   Who did Harry see Bill with?

   c.   Harry saw Bill and who?

   d.   *Who did Harry see Bill and?

SD appears to play no role whatsoever in the child's understanding of the difference observed, but auxiliary inversion happens still. To see this, consider:

(9)   Harry walks

How do we form the interrogative? The explicit rule SD is worthless: there is no auxiliary verb to invert with the subject, and *walks Harry* is nonsense (it is a perfectly good VP, of course, but it is not a sentence). In English, auxiliary verbs are simply repositories for agreement features of tense, person, number, etc., i.e., heads of IP projections. The verb *walk* is inflected to agree with the subject *Harry*. If, then, auxiliary inversion is general, we would predict the movement of the inflectional properties from the verb stem. This we find:

(10)   Does Harry walk?

Note that the verb *walk* is now in infinitive form and the inflection has risen to attach to the pleonastic verb *do* which has the features [3rd person, present, singular] that *walk* has in (9). This effect is commonly referred to as *do support*, and is witnessed in (8) with past tense *did*.

The same principle applies with *wh*-movement from complements. Consider (11):

(11)a.   Harry said that Bob will meet Mary

   b.   Who did Harry say will meet Mary?

   c.   Whom did Harry say that Bob will meet?

Prima facie, there is competition for movement here: the auxiliary *will*, *per* SD, or the inflection on the matrix verb *say*, *per* the cases above, or perhaps even both. Nevertheless, children invariably target the matrix inflection, which moves head-to-head to form head CP *did*, just as it does with the simpler cases of *wh*-movement (we shall soon see why *only* the matrix inflection can move). But this uniformity does not make the situation simpler for the empiricist child. For the child to target the main inflectional head in any given (non-monoclausal) construction necessarily involves it excluding other inflections that rightly move in other, very similar, constructions. Thus, the child must be sensitive to matrix clause, relative clause and complement clause. It makes no sense to think that a child can fixate on, or apply, a general rule without knowing the structure to which the rule applies, and here we are seeing that the structures get increasingly complicated and so the appropriate 'rule' gets reciprocally more abstract, i.e., more distant from the morphology and linear phonology of that to which the child has access in its pld. Needless to say, no rule such as 'front the first piece of inflectional morphology' will do because the subjects here can have relatives attached just like any other subject. Nor will SD give the right result, as demonstrated. Indeed, 'inversion' is witnessed where there is apparently nothing with which to invert.
Consider:

(12)a.   Knowing that Harry will steal the car bothered him

   b.   Did knowing that Harry will steal the car bother him?

   c.   Will knowing that Harry stole the car bother him?

   d.   *Will knowing that Harry steal the car bothered him?

(12)b is the interrogative form of (12)a with the inflection of *bother* rising and *do supported*. *Bother* is here the matrix verb and carries the main inflection of the sentence. As (12)d demonstrates, *will* cannot rise from complement position. (In technical terms, the auxiliary would have crossed two head projections: NP *Harry* and CP *that*. A *shorter* movement is available as witnessed in (12)b, where only head CP is crossed. SD may be viewed as a rough instance of such a general economy on head-to-head movement.) Also note that while *will* can rise as in (12)c, it doesn't rise from complement position, i.e., (12)c is not a questioning of (12)a. The *will* evidently rises from matrix position since the complement verb is inflected, and so is not questioned. (Again, technically, only the head CP is crossed). *We* know all of this, but the distinct constructions the empiricist child must generalise to – without the benefit of negative data such as (12)d – have again multiplied, and in a curious way.

SD speaks of inverting the auxiliary with the subject NP, but with (12)a there is no overt subject. Even so, inversion occurs. On its natural reading, the understood subject of (12)a is co-referential with *him*. The long accepted understanding of such constructions is that the sentence is represented with a covert (pronominal-like) subject PRO that *controls him*. The inflection, then, does indeed hop over the subject to head CP with PRO as the nominal, but in so doing it not only starts off to the left of the overt item that is the understand subject, but it then moves as far from the item as possible. Clearly, this is something a child cannot see or hear, and it appears to contradict the putative general rule.

As indicated just above and earlier on, the current explanation of these phenomena involves a notion of *locality* or *shortest movement*, but the locality at issue, as should be evident, is not linear, but essentially structural, to do with the number of head projections a permissible movement might cross in competition with other possible movements. Also, what induces the movement are quite abstract: features of tense and agreement that are variously realised in the sentence morphology. Further, elements 'appear' via do support which are not present in the declarative; they occur for structural reasons, to take on the moved inflection to form a CP head. Still further, all these phenomena are sensitive to categories that are covert. (This is to say nothing of *why* there is movement at all. The thought that such 'imperfection' is driven precisely to *check* inflectional morphology is at the centre of the minimalist program (Chomsky, 1995). Suffice it to say, the above data may be readily understood as challenges to the empiricist to find a rule which allows the child to check inflectional morphology in accord with UG constraints on interrogatives. That picture looks even worse for the empiricist.)

What is emerging is that the child needs to hypothesise a rule, as the empiricist would have it, that applies across the board to many different kinds of constructions. Cowie's claim that the empiricist child can have a preference for rules that appeal to "unobservables" is appearing to be even more hollow than it did initially. Not even SD is adequate to capture the required generalisation. But the child, we are supposed to think, can statistically induce the common properties of the variety of constructions considered without foreknowledge of syntactic structure: CP heads, IP, heads, *do support*, PRO, species of clauses, etc. This is not an a priori argument, of course, but it is sufficient to display the muddle Cowie is in when she claims that the issue is about ad hoc rule decapitation. That is precisely not what is at issue.

The empiricist might protest that the constructions so far considered, no matter their differences, are all interrogative; perhaps the child gets enough cues from the prosody and semantics of questioning (*some* mothers ask a lot of questions, after all) to discover the common pattern of inversion. This complaint would be mere rhetoric even if inversion were restricted to interrogatives, but it is not so restricted. Besides which, Cowie for one produces no evidence that an empiricist can account for SD, let alone the cluster of principles to do with shortest head-to-head movement required to cover all cases.

Consider the negative not. In most dialects of English, *not* occurs after the first auxiliary and modifies the main verb; e.g.:

(13)a.   Harry can walk

   b.   Harry can not walk

The formation of negative declaratives, then, is sensitive to the position of inflectional morphology just as we found with interrogatives. It might seem that the child could learn this easy enough and that it has nothing much to do with the interrogative case. Consider, though, sentences with no auxiliary, such as (9). How do we negate them? In English, unlike in logic, negation tends to be internal, but (14) is nonsense:

(14)   * Harry not walks

The problem with (14) is similar to that of *walks Harry*, viz.,, if we want to negate or question a sentence with no auxiliary verb, such as (9), we cannot move the main verb, but we must raise the inflectional affix. Thus:

(15)   Harry does not walk

Again, the inflectional features ([3rd person, present singular]) of the affix -s attaches to the pleonastic *do*, leaving the main verb in infinitive form. To understand negation in general, then, the child must understand the differing placements of inflectional properties just as is required in the general understanding of interrogatives.

There are many other examples of constructions which exhibit the same pattern, but the idea should now be clear. The moral is that the more disparate the data is to which a principle applies, the more abstract the knowledge or generalisation is that the child must grasp. Cowie does think that the empiricist can glean 'deep regularities', not just surfaces ones. I know of no proof that this is impossible, but we are supposed to be thinking about science, not metaphysics; the appeal to deep regularities is simply a priori wishful thinking, for such supposed available regularities amount to no more than those which *must* obtain if empiricism is not to be false. Thus, without even a hint of what such regularities are, Cowie's claim reduces to the true but trite thought that empiricism is not necessarily false. The only possible reason for thinking this interesting is that 'Chomskyans' think empiricism logically false (see Cowie, 2001, 240). But who are these 'Chomskyans'?

In short, what we find when we investigate a single 'abstract rule' is that its content is a complex function of a deep set of conditions that are not mirrored at the surface. It is such a realisation that led to the P&P approach, and the attempt to explain evident patterns as the outcome of the interaction of a number of quite simple principles (the above discussion could be easily substituted for one concerning anaphoric dependence, or empty categories, or case, etc.; and this is to say nothing about the similarities across languages). I have not sought to explain why English exhibits this pattern of inflectional movement between types of interrogatives and negations; there is no settled answer, and to explore the various avenues is not only beyond my present scope, but also not to the point. The linguist does not produce rules to account for given constructions, but attempts to discern an underlying structure of which given constructions are partial reflections: there is no Hydra at which an empiricist might swing her sword. Since Cowie does not doubt the correctness of the 'rules', one is left to conclude that the empiricist is beholden to explain a complex pattern of similarity and variation; there is no benefit to guessing that a given construction could be accounted for by such and such statistical analysis without concomitant light being shed on other constructions.

These considerations do not amount to a knock-out blow, but that is not their intention. It is because Cowie so badly misconstrues both the purpose and structure of the contemporary generative program that empiricism

seems to her to be so immune to the familiar objections. When Chomsky (e.g., 1975, 1980, 1991) challenges the empiricist to produce "substantive proposals", he is not attempting to refute empiricism; he is simply asking: Where is the alternative account of how the child may acquire the complex structures revealed by current linguistics? This is not rhetoric, but normal scientific inquiry. Cowie (1999, 272) scolds the nativist for 'I'm the only president you've got'-style arguments; this is quite jejune. No-one seriously involved in linguistics and related disciplines is trying to gain any knock-outs. It is because empiricism and its behaviourist progeny so grievously underestimated the complexity of what a speaker knows that it is apposite to demand concrete proposals which are sensitive to the many data. Otherwise, there is 'nothing to discuss'. Chomsky is not trying to win by default, he is just not concerned with idle a priori possibilities.

4. CONCLUDING REMARKS

There is much in Cowie's critique, and much more, of course, in the wider literature, that I have not even hinted at. My brief has only been to defend linguistic nativism against Cowie's central deflationary criticisms. In so doing, Cowie is revealed, I think, as the latest in a long wearisome line of philosophers who have sought to challenge the assumptions of the generative program by showing that they are not necessarily true. Chomsky's (1968/72, 1975, 1986, 2000) response to such critics has remained constant: generative linguistics is an empirical research program that carries no a priori assumptions; the only worthwhile way to assess its validity is to do the science and wait and see. In particular, proposing vague alternatives with no theoretical or empirical support does not constitute an interesting scepticism. Cowie, for sure, is at the enlightened end of the spectrum of Chomsky's assailants. Even so, ultimately, her criticism amounts to the claim that the perceived complacency of the 'Chomskyan' is not appropriate. Quite! But this is a perception; the business end of linguistics is where our attention should be directed, not at the polemics, which have distracted Cowie to foisting false claims upon Chomsky among many others.

NOTES

1. Chomsky's notion of a faculty is an *intentional* one in that it concerns what a speaker *knows*, it does not concern how a speaker possesses such knowledge. Even so, Chomsky clearly understands such knowledge to be the content of a dedicated device, which we may call the language faculty, for we are individuating the module in terms of its content. On this issue at least, I side with Cowie against Fodor (2001, §2.1).

2. My work for this paper was completed before I read either Fodor or Matthews on Cowie. Happily, however, it may be read as a continuation and extension of some of the points raised in Matthews' review. This, if I may say so, is not surprising, for the lie of the empirical land, *pace* Cowie, is (or should be) untendentious.

3. The notion that the language faculty is interfaced with just articulatory and conceptual systems is a simplifying assumption. For example, deaf children acquire sign with the same alacrity as their speaking counterparts acquire normal speech (Klima and Bellugi, 1979). More tendentiously, I have suggested a ToM module-language faculty interface (Collins, 2000).

4. Cowie's Chapter 10 is largely given over to questioning the P&P approach. She is perfectly correct in her judgement that vagueness and ambiguity surround what parameters there are, what their initial settings are, and what triggers are needed. But none of this is news to anyone in linguistics. Moreover, P&P has always been a working assumption rather than any kind of a priori stipulation; Chomsky (1995, 7) dubs it "a bold speculation rather than a particular hypothesis". While Cowie (1999, 174) contends that such modest statements are the "rare exception", she fails to produce a single quotation to support a stronger reading, and fails to mention the ubiquitous passages to the contrary.

5. See Pinker (1999) for a survey of the recent data and theory.

6. See Newmeyer (1991) for an excellent discussion of the fluctuating status of 'rules' within the history of the generative tradition.

7. Within the minimalist program (Chomsky, 1995), even grammatical categories are treated as taxonomic epiphenomena.

REFERENCES

Bohanon, J., B. MacWhinney, and C. Snow: 1990, 'No Negative Evidence Revisited: Beyond Learnability or Who Has to Prove What to Whom', *Developmental Psychology* **26**, 221–226.

Bohanon, J. and L. Stanowicz: 1988, 'The Issue of Negative Evidence: Adult Responses to Children's Language Errors', *Developmental Psychology* **24**, 684–689.

Brown, R. and C. Hanlon: 1970, 'Derivational Complexity and Order of Acquisition in Child Speech', in J. Hayes (ed.), *Cognition and the Development of Language*, John Wiley, New York.

Chomsky, N.: 1957, *Syntactic Structures*, Mouton, The Hague.

Chomsky, N.: 1965, *Aspects on the Theory of Syntax*, MIT Press, Cambridge, MA.

Chomsky, N.: 1968/72, *Language and Mind* (enlarged edition), Harcourt Brace Jovanovich, New York.

Chomsky, N.: 1975, *Reflections on Language*, Fontana, London.

Chomsky, N.: 1980, *Rules and Representations*, Columbia University Press, New York.

Chomsky, N.: 1986, *Knowledge of Language: Its Nature, Origin and Use*, Praeger, Westport.

Chomsky, N.: 1991, 'Linguistics, a Personal View', in A. Kasher (ed.), *The Chomskyan Turn*, Blackwell, Oxford.

Chomsky, N.: 1995, *The Minimalist Program*, MIT Press, Cambridge, MA.

Chomsky, N.: 2000, *New Horizons in the Study of Language and Mind*, Cambridge University Press, Cambridge.

Collins, J.: 2000, 'Theory of Mind, Logical Form, and Eliminativism', *Philosophical Psychology* **13**, 465–490.

Cowie, F.: 1999, *What's Within? Nativism Reconsidered*, Oxford University Press, Oxford.

Cowie, F.: 2001, 'On Cussing in Church: In Defence of *What's Within*', *Mind and Language* **16**, 231–245.

Crain, S.: 1991, 'Language Acquisition in the Absence of Experience', *Brain and Behavioural Sciences* **14**, 597–615.

Crain, S. and R. Thornton: 1998, *Investigation in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. MIT Press, Cambridge, MA.

Demetras, M., K. Post, and C. Snow: 1986, 'Feedback to First Language Learners: The Role of Repetitions and Clarification Questions', *Journal of Child Language* **13**, 275–292.

Feldman, H., S. Goldin-Meadow, and L. Gleitman: 1978, 'Beyond Herodotus: The Creation of Language by Linguistically Deprived Deaf Children', in A. Lock (ed.), *Action, Symbol, and Gesture*, Academic Press, New York.

Fodor, J.: 1998: *Concepts: Where Cognitive Science Went Wrong*, Oxford University Press, Oxford.

Fodor, J.: 2001: 'Doing Without *What's Within*: Fiona Cowie's Critique of Nativism', *Mind* **110**, 99–148.

Gleitman, L.: 1990, 'The Structural Sources of Word Meaning', *Language Acquisition* **1**, 33–55.

Gleitman, L. and E. Warner: 1982, 'Language Acquisition: the State of the Art', in E. Warner and L. Gleitman (eds.), *Language Acquisition: The State of the Art*, Cambridge University Press, Cambridge.

Gordon, P.: 1990, 'Learnability and Feedback', *Developmental Psychology* **26**, 217–220.

Grimshaw, J.: 1981, 'Form, Function, and the Language Acquisition Device', in C. Baker and J. McCarthy (eds.), *The Logical Problem of Language Acquisition*, MIT Press, Cambridge, MA.

Grimshaw, J. and S. Pinker: 1989, 'Positive and Negative Evidence in Language Acquisition', *Behavioural and Brain Sciences* **12**, 341.

Heath, S.: 1983, *Ways with Words: Language, Life, and Work in Communities and Classrooms*, Cambridge University Press. Cambridge.

Hirsh-Pasek, K., R. Treiman, and M. Schneiderman: 1984, 'Brown and Hanlon Revisited: Mother's Sensitivity to Ungrammatical Forms', *Journal of Child Language* **11**, 81–88.

Klima, E. and U. Bellugi: 1979, *Signs of Language*, Harvard University Press, Cambridge, MA.

Lightfoot, D.: 1993, *How to Set Parameters: Arguments from Language Change*, MIT Press, Cambridge, MA.

Maratsos, A.: 1982, 'The Child's Construction of Grammatical Categories', in E. Warner and L. Gleitman (eds.), *Language Acquisition: The State of the Art*, Cambridge University Press, Cambridge.

Marcus, G.: 1993, 'Negative Evidence in Language Acquisition', *Cognition* **46**, 53–85.

Matthews, R.: 2001, 'Cowie's Anti-nativism', *Mind and Language* **16**, 215–230.

Moerk, E.: 1991, 'Positive Evidence for Negative Evidence', *First Language* **11**, 219–251.

Morgan, J. and L. Travis: 1989, 'Limits on Negative Information in Language Input', *Journal of Child Language* **16**, 531–552.

Nakayama, M.: 1987, 'Performance Factors in Subject-Auxiliary Inversion by Children', *Journal of Child Language* **14**, 113–125.

Newmeyer, F.: 1991, 'Rules and Principles in the Historical Development of Generative Syntax', in A. Kasher (ed.), *The Chomskyan Turn*, Blackwell, Oxford.

Newport, E., H. Gleitman, and E. Gleitman: 1977, 'Mother, I'd Rather Do It Myself: Some Effects and Non-effects of Maternal Speech Style', in C. Snow and C. Ferguson (eds.), *Talking to Children: Language Input and Acquisition*, Cambridge University Press, Cambridge.

Newport, E. and R. Meier: 1985, 'The Acquisition of American Sign Language', in D. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition*, Erlbaum, Hillsdale, NJ.

Penner, S.: 1987, 'Parental Responses to Grammatical and Ungrammatical Child Utterances', *Child Development* **58**, 376–384.

Pinker, S.: 1984, *Language Learnability and Language Development*, Harvard University Press, Cambridge, MA.

Pinker, S.: 1987, 'The Bootstrapping Problem in Language Acquisition', in B. MacWhinney (ed.), *Mechanisms of Language Acquisition*, Erlbaum, Hillsdale, NJ.

Pinker, S.: 1999, *Words and Rules: The Ingredients of Language*, Weidenfeld and Nicolson, London.

Pullum, G.: 1996, 'Learnability, Hyperlearning, and the Poverty of Stimulus', in J. Johnson, M. Junge, and J. Moxley (eds.), *Proceedings of the 22nd Annual Meeting: General Session and Parasession on the Role of Learnability in Grammatical Theory*, Berkeley Linguistics Society, Berkeley.

Putnam, H.: 1971, 'The "Innateness Hypothesis" and Explanatory Models in Linguistics', in J. Searle (ed.), *The Philosophy of Language*, Oxford University Press, Oxford.

Saffran, J., R. Aslin, and E. Newport: 1996, 'Statistical Learning by 8-Month Old Infants', *Science* **274**, 1926–1928.

Sampson, G.: 1989, 'Language Acquisition: Growth or Learning?' *Philosophical Papers* **18**, 203–240.

Schieffelin, B. and A. Eisenberg: 1981, 'Cultural Variation in Children's Conversations', in R. Schiefelbusch and D. Bricker (eds.), *Early Language: Acquisition and Intervention*, University Park Press, Baltimore.

J. Collins
Philosophy, SOC
University of East Anglia (Norwich)
Norwich, NR4 7TJ
U.K.
E-mail: jcollins42@compuserve.com