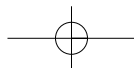
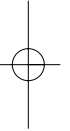
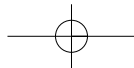
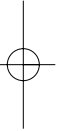
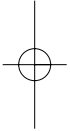
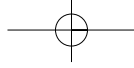


# HOW MUCH KNOWLEDGE OF LANGUAGE IS INNATE?





**CHAPTER  
FOUR****Irrational Nativist Exuberance***Barbara C. Scholz and Geoffrey K. Pullum*

The protracted dispute over the degree of independence of language acquisition from sensory experience often degenerates into an unsavory cavalcade of exaggerated claims, tendentious rhetoric, and absurd parodies of opposing views.<sup>1</sup> In this chapter we try to distinguish between partisan polemics and research programs. If either side of the partisan dispute about the acquisition of syntax were as stupid as the opposing side alleges, the free-for-all would not be worthy of serious attention; but in fact we think there are two important complementary research programs for syntax acquisition involved here.

We are skeptical about recent triumphalist claims for linguistic nativism, and this may lead to us being mistaken for defenders of some sort of “empiricism.”<sup>2</sup> But touting empiricist stock is not our project. Curbing the excesses of irrational nativist exuberance is more like it. We argue that it is premature to celebrate nativist victory (as Laurence and Margolis, 2001, seem to be doing, for instance),<sup>3</sup> for at least two reasons. First, the partisan dispute is too ill-delineated to reach a resolution at all, because of a persistent tendency to conflate non-nativism with reductive empiricism, and because of equivocations rooted in the polysemy of the word “innate” (section 2). And second, linguistic nativist research programs need theories of learning – exactly what non-nativist research programs aim to develop (section 3).

Although we use “linguistic nativism” throughout this paper to denote a number of contemporary views about the acquisition of syntax, the reader will note that we tend to avoid using “innate.” We try instead to address the specifics that the term “innate” often seems to occlude rather than illumine: the extent to which the acquisition of syntax proceeds independently of the senses, for example, and the extent to which it depends on generally applicable human cognitive capacities. The traditional “empiricist” claim is that the syntactic structure of languages (like everything else) is learned from sensory input. This could be false in at least two ways: it could be that the syntactic aspects of language are not acquired at all, but are antecedently

known or available; or it could be that the influence of sensory input is not anything that could be described as “learning” in any normal sense. This distinction will be important below when we consider a debate about whether the products of “triggering” count as being innate.

Linguistic nativists refer to the constraints they posit on language acquisition as “linguistic knowledge” or “Universal Grammar” (abbreviated “UG”). We are going to assume they mean the following conjunction: (i) language acquisition is constrained by either *universal linguistic principles* or *biases due to language-specialized cognitive mechanisms*, and (ii) these are either *unacquired* or *acquired but not learned*. The non-nativist, by contrast, denies the conjunction of (i) and (ii).

Notice that the non-nativist’s rejection of unacquired but language-specialized cognitive mechanisms does not imply a rejection of unacquired *non-cognitive* mechanisms (e.g., perceptual ones) that constrain language acquisition. And of course the non-nativist view is compatible with general cognitive mechanisms acquiring language-specialized functions over (ontogenetic) time. Disputes about syntax acquisition between contemporary linguistic nativists and their opponents concern *what* constrains or biases language acquisition, not whether it is constrained or biased. Language acquisition is, for example, unquestionably constrained by quantitative resource bounds and sensory channel limits.

## 1 Contemporary Non-Nativism and Historical Empiricism

The classic empiricist slogan states that there is “nothing in the intellect which was not previously in the senses” (Aquinas, *Summa Theologica*, Ia). Nativism is often taken to be the negation of empiricism: the view that at least one thing is in the intellect that was *not* acquired from the senses. But this is too weak to be an interesting form of contemporary nativism. It would surely be a pyrrhic victory if linguistic nativism were true simply in virtue of one solitary unacquired or unlearned contentful linguistic principle, everything else being learned.<sup>4</sup> And it would make it a mystery why nativist linguists have attempted to establish the existence of so many such principles, and have emphasized their abundance.<sup>5</sup> For the purposes of this chapter, we take linguistic nativism to be the view stated in (1):

- 1 Most of the acquisition of natural languages by human beings depends on unacquired (or acquired but unlearned) linguistic knowledge or language-specialized cognitive mechanisms.

This psychological generalization quantifies over unacquired (or acquired but unlearned) knowledge and mechanisms specialized for language.<sup>6</sup> The research program of linguistic nativism aims to show, proposition by proposition and mechanism by mechanism, that very little knowledge of syntactic structure is acquired or learned from sensory stimuli. Thus the discovery of one (or even a few) language-specialized cognitive mechanisms does not resolve the partisan nativist/non-nativist dispute. Even after the discovery of one genuinely unacquired linguistic principle, the continued

development of both nativist and non-nativist research programs would and should continue.

*Non-nativism* with regard to language acquisition is the view stated in (2):

- 2 Very little of the acquisition of natural languages by human beings depends on unacquired (or acquired but unlearned) linguistic knowledge or language-specialized cognitive mechanisms.

This too is a psychological generalization about how particular languages are acquired, and it competes with (1). Advocates of (2) advance several distinct but related research programs in language acquisition that include stochastic, constructivist, and connectionist learning theories.

What exactly is at issue between linguistic nativists and non-nativists depends, in part, on identifying what being “innate” amounts to (which we turn to in section 2), and how the posited unacquired language-specialized mechanisms solve certain problems of language acquisition (which we address in section 3). But one thing is clear: contemporary non-nativism is not restricted to the vaguely delineated constellation of doctrines supposedly held by the likes of John Locke in the seventeenth century and B. F. Skinner in the twentieth. Non-nativists are not obliged to defend either Locke or Skinner as having proposed a productive research program on language acquisition (see Trout, 1998, on this point). Neither Lockean reductive empiricism nor Skinnerian behaviorism is entailed by the rejection of linguistic nativism. Unacquired sensory mechanisms (perhaps specialized for speech) and general computational mechanisms are, for the non-nativist, mechanisms of learning readiness. It is an error to assume that all linguistic non-nativism is a form of reductive empiricism.

Some nativists do commit this error. Fodor (1981) provides a particularly clear example. In the context of discussing both historical empiricist and contemporary non-nativist views of lexical concept acquisition, Fodor writes:

[T]he Empiricist model says that there are two sorts of causal processes at play in concept acquisition: in the case of primitive concepts there are, one might say, brute-causal processes. The structure of the sensorium is such that certain inputs trigger the availability of certain concepts. Punkt. Whereas in the case of complex concepts, there are, as one might say, rational causal processes: the psychological mechanisms which mediate the availability of such concepts realize an inductive logic, and the experiences of the organism stand in a confirmation relation to the concepts whose availability they occasion. (Fodor, 1981, p. 273)

Fodor’s “primitive concepts” correspond to the simple, unstructured ideas of British empiricism. Though activated or triggered by sensory experience, their acquisition is not explained by any psychological learning theory. Fodor alleges that all empiricist theories of concept acquisition are *reductive*: all learned complex concepts are reducible to triggered primitive concepts, and for the complex ones, “concept learning involves the inductive fixation of certain beliefs [complex concepts] and the mechanisms of concept learning are realizations of some species of inductive logic” (1981, p. 267).

Given this reductive characterization, Fodor struggles to specify the point of disagreement between nativists and both historical empiricists and contemporary non-nativists (1981, pp. 279–83). Do they disagree “over which concepts are primitive”? Do nativists deny “that the primitive concepts constitute an epistemologically interesting set”? Do contemporary empiricists accept that “the primitive concepts are the ones whose attainment I can’t eventually account for by appeal to the mechanisms of concept learning”? Unsurprisingly, his effort to locate the precise point of difference fails. A dispute that is purportedly about the resources required for language acquisition is miscast as a dispute about the acquisition of unstructured concepts and the failure of reductive empiricism, all wrapped up in an empiricist theory of justification rather than a theory of concept acquisition.

Contemporary non-nativist psychology need not be either atomistic or reductive. Consider for example the non-reductive conjectures about the sense-based acquisition of natural-kind concepts developed by Boyd (1981, 1991) and Kornblith (1993), crucial to their understanding of scientific realism. Boyd finds in certain (unofficial) Lockean views the suggestion that natural kinds are primitive, complex, structured, homeostatic clusters of properties, and our sense-based concepts of them are complex homeostatic cluster concepts. What Boyd seems to reject is that primitive concepts are non-complex, and that non-complex concepts are epistemologically interesting. We will not extend these ideas to language acquisition here, but we note that Christiansen and Curtin (1999) appear to be applying them to word individuation.

Fodor, however, is certainly right about at least two things. First, any coherent research program in language acquisition must accept that some acquisition mechanisms are not acquired. All parties must concede this on pain of a vicious regress of acquired mechanisms for acquisition (see Block, 1981, p. 280). But Chomsky presses this point to a parodic extreme:

To say that “language is not innate” is to say that there is no difference between my granddaughter, a rock, and a rabbit. In other words, if you take a rock, a rabbit, and my granddaughter and put them in a community where people are talking English, they’ll all learn English. If people believe that, then they’ll believe language is not innate. If they believe that there is a difference between my granddaughter, a rabbit, and a rock, then they believe that language is innate. (Chomsky, 2000, p. 50)

The very possibility of a non-nativist research program is trivialized by saddling it with the view (which no one holds) that humans have the same language acquisition mechanisms as rocks. If all the alternatives to nativism are malignly depicted as inherently dumb, then the triumph of nativism will be hollow.

The second thing Fodor is right about is that creatures who acquire *symbolic* representations or concepts by means of hypothesis formation and testing must antecedently have some concepts in which to frame hypotheses. But this is a claim about a specific kind of a prioristic learning theory and symbolic representation-alism, not the distributed representations and stochastic learning theories investigated by contemporary non-nativist research programs.

It is also an error to understand non-nativism as claiming that what is learned or acquired by stochastic means can only be strictly proportional to the stimulus – as

if what is so acquired is never anything more than an exact copy of the statistical distributions in the stimulus. This error is frequently made by linguistic nativists. For example, Lidz et al. (2003) write that “It is hopeless to suppose that learning is responsive (solely) to input frequency, because the first word [that children acquire] in English vocabulary is not *the*.” As Elman (2003) notes, it is an error to take stochastic learning theory to hypothesize that children learn statistics, i.e., they merely copy or memorize stimulus frequency patterns. On the contrary, stochastic learning theory holds that language learning is based on complex, higher-order properties of stochastic patterns in sensory experience, not a mere tabulation of frequency of patterns. To take children’s (or adults’) sense-based stochastic acquisition abilities to be limited to frequency detection and tabulation greatly underestimates their power. One leading question in statistical language acquisition research concerns the kinds of stochastic patterns infants can acquire (Saffran et al., 1996).

Nativists also sometimes mistakenly assume that the only kind of linguistic stimulus that could be relevant to language acquisition is the presence or absence of certain individual strings in the primary linguistic data. The assumption that rare or even absent strings would have to occur frequently for stochastically based learning to succeed oversimplifies (without investigation) the relevant distributional properties of the data (Lewis and Elman, 2001; Elman, 2003). Reali and Christiansen (forthcoming) provide further evidence that the relevant features of the stimulus for statistically based language acquisition models are the stochastic properties of the overall input, not just the presence or absence of individual strings therein. And see also Saffran et al. (1996), Aslin et al. (1998), Gomez (2002), and Saffran and Wilson (2003) for evidence that children are effective statistical learners.

Non-nativist researchers on language acquisition are free to either accept or reject historical empiricist doctrines, because contemporary linguistic non-nativism is not a form of reductive empiricism. It is merely a rejection of (1).

## 2 What Innateness Is

The hypothesis that some features of natural languages are acquired by triggering is characteristic of the “principles and parameters” theory.<sup>7</sup> Parameters are supposed to “reduce the difficulty of the learning problem” (Gibson and Wexler, 1994, p. 407). Parametrized universal principles are hypothesized to facilitate language acquisition by reducing what must be learned from sensory experience about the systematic parochial variations of natural languages.<sup>8</sup> A parameter does not specify a single property common to all natural languages. Rather, it specifies a fixed set of mutually exclusive linguistic properties, of which any given natural language can have exactly one.

Parameters are alleged to be unacquired. What is acquired is a particular setting of a parameter, by the process of being triggered by an environmental stimulus or range of stimuli. For example, “initial” might be one possible setting for a parameter governing position of lexical head (e.g., the verb in a verb phrase), and “final” the other setting, ruling out the possibility of any language in which lexical heads are positioned, say, as close to the middle of a phrase as possible. The debated issue

in the philosophy of linguistics literature is whether the *products* of triggering processes (parameters set to particular values) count as being innate.

On one side of the debate, Samuels (2002), developing Fodor's view, claims that a cognitive structure (e.g., a parameter or a concept) is innate if and only if it is *psychologically primitive*. For Samuels a structure is psychologically primitive when it is "posited by some correct scientific psychological theory" but its acquisition has *no* explanation in that theory (Samuels, 2002, p. 246). A primitive process is "a 'brute causal' process that is not mediated by any psychological states," e.g., not mediated by information uptake or processing. Samuels takes triggering to be a paradigmatically primitive psychological process (p. 247), since he thinks any explanation of triggering will be biological, not psychological. He also seems to think that the products of triggering processes are themselves psychologically primitive.

On the other hand, Arieu (1996, 1999, 2003), following Waddington (1940), has argued that innateness is *environmental canalization*, which entails that "development is *insensitive* to particular environmental conditions." This is the idea that what is innate is developmentally rigid and hard to change. Any trait that is environmentally canalized is "buffered against environmental perturbation" (Arieu, 2003). Processes like triggering and their products are sensitive to, and depend on, particular environmental cues. Thus they are *not* innate in Arieu's sense. However, on his view, traits acquired by triggering are not learned either. Whatever is acquired by triggering is the result of a *sui generis* developmental process that is sensitive to environmental stimuli but is neither innate nor learned.

Samuels stipulates that whatever is innate is not learned. Arieu replaces that traditional dichotomy with a trichotomy: *innate* versus *triggered* versus *learned*. Our interest in this dispute is not in taking sides, but in pointing out that Samuels and Arieu are probably both right about triggering because both "innateness" and "triggering" are used polysemously in the linguistics literature.

The ambiguous use of "triggering" was noticed 30 years ago by Stich (1975). At least two undeveloped metaphors of acquisition by triggering are found in the literature. One is based on Plato's parable of the slave boy in the *Meno*. Stich writes:

a trigger or a catalyst . . . sets off a process that results in the acquisition of the belief. But, as a catalyst it is not part of the end product of the chemical reaction, so the questioning process [of the slave boy by Socrates] does not supply the content of the belief . . . the content of the belief was contained in the boy. (Stich, 1975, p. 14)

The antecedently given range of possible parameter settings are "in the slave boy." The information in the activated parameter is *not* acquired by reasoning or inference from information in the environmental trigger. It is inherent in the boy's pre-existing parameter.

In what follows we use *instantaneous acquisition* for this kind of parameter setting by (Fodorian) triggering. In instantaneous acquisition no information in the environmental trigger informs or is taken up into the product of the triggering process: e.g., exposure to ambient temperatures of above 90°F might cause the head parameter to be set to strict verb-final clause structure.



Under the second triggering metaphor (Stich, 1975, p. 15), the parameter is not merely set off or activated. Rather, the *information* in the environmental trigger is relevant to the information content of the product of the triggering process (though the latter is not inferred from the former). As Gibson and Wexler (1994, p. 408) characterize it, triggers are “sentences in the child’s experience that point directly to the correct settings of parameters”; indeed, for any setting of a parameter “there is a sentence that is grammatical under that setting but not under any other.” Exposure to a trigger “allows the learner to determine that the appropriate parameter setting is the one that allows for the grammaticality of the sentence.” Gibson and Wexler go on to develop this view (see their “Triggering Learning Algorithm,” 1994, pp. 409–10): if a trigger fails to be grammatical as analyzed by the currently entertained grammar, the learning algorithm modifies a parameter setting to see if that will permit the trigger to be analyzed successfully (and changes it back again if not).

Henceforth we use the term *accelerated acquisition* for this kind of informationally triggered parameter setting.

By claiming that the learner determines the setting, Gibson and Wexler mean that the uptake of linguistic information in the trigger is necessary for the acquisition of a particular setting for a parameter. The familiar expression “poverty of the stimulus” picks up on impoverished information in the triggering stimulus by contrast with the richness of the information in the set parameter in both instantaneous and accelerated acquisition. But the two triggering processes are distinct. We speculate that an over-emphasis on arguments from the poverty of the stimulus has led many philosophers and linguists to overlook the differences between these two kinds of triggering.

However, Gibson and Wexler do distinguish sharply between their concept of triggering and Fodorian triggering (instantaneous acquisition) which is “supposed to mean something like an experience that has nothing theoretically to do with a parameter setting, but nevertheless determines the setting of a parameter” (Gibson and Wexler, 1994, p. 408, n. 2).

Instantaneous acquisition is a brute-causal psychological process, unmediated by intervening psychological states, exactly as Samuels says. Thus the product of instantaneous acquisition is innate for Samuels. For Gibson and Wexler, on the other hand, discrimination and uptake of the *information* in the stimulus mediates the setting of a parameter. Parameter setting by accelerated acquisition is neither psychologically primitive nor environmentally canalized, so its products are not innate in either Samuels’ or Arieuw’s sense.

Doubtless Gibson and Wexler would reject both Samuels’ and Arieuw’s concepts of innateness because they claim the product of their preferred triggering process is innate, although neither psychologically primitive nor canalized. But if it is, then there must be some other concept of innateness to vindicate their claim. Quite a few are on offer.

Sober (1998) has argued that all that is left of the pre-scientific concept of innateness is the idea that what is innate is *invariant* across environments. But it is immediately clear that invariance innateness won’t do: triggering is supposed to explain the acquisition of linguistic structures that systematically vary across natural languages. Of course, parameters that have not yet been set are in a sense invariant. But they do not explain how infants acquire knowledge of parochial aspects of their languages.

Some scientists talk as if what is *universal* across all typical members of the species, or across all natural languages, is innate (see, e.g., Barkow et al., 1992); but, *ex hypothesi*, the products of triggering are not universal.

Stich (1975, p. 12) considers a Cartesian *dispositional* analysis of innate beliefs: a belief is innate for a person just in case “that person is disposed to acquire it under any circumstances sufficient for the acquisition of any belief.” But this lends no support to any advocate of the idea that the products of triggering are innate. Knowledge of particular languages that is acquired by the triggering of a parameter requires special circumstances.

Gibson and Wexler should probably not consider reverting to the idea that what is innate is *known a priori*. First, if a priori acquisition is defined as “what is acquired independently of any specific experience,” then the products of instantaneous and accelerated acquisition, which depend on the experience of specific sensory triggers, are not innate; and defining a priori knowledge as “what is known on the basis of reason alone” fails because the products of all triggering processes are, by definition, not acquired by means of inference or reason.

Bealer (1999) has more recently articulated a concept of the a priori through evidence that “is not imparted through experiences but rather through intuitions” (p. 245). According to Bealer, “For you to have an intuition that A is just for it to *seem* to you that A” (p. 247). Thus, we might say that

a trait or structure, A, is innate for S just in case S’s evidence for A is *a priori*, i.e., it cognitively, consciously, and reliably seems to S that A.

But the triggering process is, *ex hypothesi*, not consciously accessible. And the idea that a trait is innate just in case it is due to our “biological endowment” fails, since even a behaviorist thinks association mechanisms are part of our biology.

At least three of the senses of “X is innate” in the linguistics literature that we have discussed here are empirically dissociated: (i) X is a psychological primitive, (ii) X is canalized, and (iii) X is universal across all natural languages. We have also seen that linguistic nativism hypothesizes at least three distinct specialized mechanisms of language acquisition that correspond to each of these kinds of innateness: instantaneous acquisition, accelerated acquisition, and unacquired universal principles. Our point is not that any one of these conceptions of innateness is somehow illegitimate, or that one is to be preferred to the others. So far as we can tell each of these kinds of innateness could plausibly play a role in the explanation of language acquisition. Rather, our worry is that treating empirically dissociated mechanisms with the single label “innate” only obscures the detailed and accurate understanding of language acquisition that is the goal of cognitive psychology.

We are certainly not the first to notice that the blanket labeling of distinct developmental trajectories as “innate” (or “learned”) impedes scientific understanding. (Bateson, 1991; 2004, pp. 37–9) has identified seven ways in which “instinct” and “innate” are polysemous in the behavioral ecology literature, and notes that few traits are innate in all seven senses. Griffiths (2002) argues that “innateness” is undesirable as a theoretical term, since it confuses exactly what needs to be clarified. We join Bateson, Griffiths, and others in recommending that the term “innate” be abandoned in theorizing about language acquisition, because it impedes the study of language acquisition.

This recommendation is not anti-nativist. Marler (1999) posits highly specialized innate abilities, but nonetheless advocates abandoning the use of the terms “innate” and “learned”:

As all students of behavior are aware, arguments about where the emphasis should be placed in discussions of the *innate-learned* dichotomy are notoriously contentious, often becoming so strident that their scientific productivity is questionable . . . Thus one drawback to the *learned-innate* dichotomy is that it tends to polarize our thinking about how a particular behavior develops. Perhaps most importantly, it encourages us to underestimate genetic contributions to the development of *learned* behaviors. (Marler, 1999, p. 311)

Marler is not advocating some bland form of interactionism. Rather, he is suggesting that the innate-learned dichotomy encourages the “black-boxing” of significant genetic factors in way that undermines our understanding of birdsong acquisition.

If this recommendation is followed, then partisan nativist/non-nativist disputes will be less likely to be polarized. And they might be more productive if they focused on a detailed examination of *all* the mechanisms that play a role in language acquisition.

### 3 Unacquired Linguistic Universals: What Are They Good For?

Linguistic nativists have repeatedly emphasized that they think that the human infant *must* be in possession of unacquired linguistic universals (which we will henceforth refer to as ULUs). The following remarks of Hauser, Chomsky, and Fitch are representative:

No known “general learning mechanism” can acquire a natural language solely on the basis of positive or negative evidence, and the prospects for finding any such domain-independent device seem rather dim. The difficulty of this problem leads to the hypothesis that whatever system is responsible must be biased or constrained in certain ways. Such constraints have historically been termed “innate predispositions,” with those underlying language referred to as “universal grammar.” Although these particular terms have been forcibly rejected by many researchers, and the nature of the particular constraints on human (or animal) learning mechanisms is currently unresolved, the existence of some such constraints cannot be seriously doubted. (Hauser et al., 2002, p. 1577)

This recapitulates earlier claims frequently repeated and endorsed by philosophers sympathetic to linguistic nativism. Thus Lipton (1991) characterizes the linguistic nativist as holding that “children must be born with strong linguistic rules or principles that further restrict the class of languages they will learn [and are not] peculiar to a particular human language” (Lipton, 1998, p. 413), stressing the necessity of ULUs. And Laurence and Margolis (2001, p. 221) claim that “an empiricist learner . . . wouldn’t have any innate domain-specific knowledge or biases to guide her learning and, in particular, wouldn’t have any innate *language-specific* knowledge or biases.”

But from the claim that language acquisition must be affected by some sorts of bias or constraint it does not follow that those biases or constraints must stem from either ULUs or parameters. A non-nativist can readily accept biases or constraints stemming from sensory mechanisms that are specific to language but non-cognitive, or cognitive-computational mechanisms that are not language-specialized.

What tempts the defenders of nativism to believe otherwise? The matter is complex. In brief, we see three factors conspiring to tempt nativists into thinking that only ULUs could guide language acquisition: (i) an inappropriately selective skepticism based on Humean underdetermination; (ii) a highly selective faith in lexical learning by hypothesis formation and testing; and (iii) a failure to appreciate the breadth of scope of the important mathematical results set out by E. Mark Gold (1967).

The idea of studying learning by investigating the limits of an abstract pattern-learning machine originates with Ray Solomonoff in work done at the end of the 1950s (Li and Vitányi, 1997, pp. 86–92 provides a very useful history with references). Independently, it would seem, Hilary Putnam (1963a, 1963b) provided a basic impossibility result about what a machine for inductive learning could in principle accomplish: there can never be a “perfect” learning machine, because for any proposed such machine we can define a regularity that it cannot induce. Putnam’s proof strategy was later used by Gold (again, independently, it would appear) to prove a key result in the narrower domain of language learning by guessing grammars (Gold, 1967).

Gold’s enormously influential paper stimulated the development of a whole subfield of mathematical work on learning based on recursive function theory (see Jain et al., 1999, for a comprehensive survey). Conceptualizing language learning as a process of guessing a generative grammar (or, in a separate series of results, guessing a parser), Gold advocated investigation of “the limiting behavior of the guesses as successively larger and larger bodies of information are considered” (Gold, 1967, p. 465). He obtained both pessimistic and optimistic results. On one hand, he showed that there was a sense in which for all interesting classes of generative grammars<sup>9</sup> the learning problem was unsolvable, because what has to be learned is deductively underdetermined by the positive evidence (evidence about what *is* in the language; successful learning from such evidence is called “identification in the limit from text”). On the other hand, he showed that if the evidence is an information sequence covering both *what is* and *what is not* in the language, the learning problem is solvable for a huge range of classes of languages.

The pessimistic results depend on a number of assumptions. We summarize those relevant to “identification in the limit from text” in (3).

- 3 (a) The input to the learning procedure consists of strings.
- (b) The strings presented are all from the target language (no evidence about what is not in the language is provided).
- (c) Success is defined as reaching, in the limit, a unique correct generative grammar for the target language after a finite number of inputs, and sticking with it in the face of all further data.
- (d) The learning procedure selects a generative grammar to be the current hypothesis, testing it against the current datum, abandoning it permanently if it is incompatible with that datum, and sticking with it otherwise.

Since human children do learn languages, and Gold has apparently proved that they can't, we face a paradox. The only plausible response is to reject one or more of the assumptions leading to it. That is, one or more of Gold's assumptions must not hold for child language learners (see Scholz, 2004, for further discussion of this point). Many contemporary linguistic nativists respond by rejecting one them, namely (3d), the assumption that language acquisition proceeds by hypothesis formation and testing. The positive alternative they propose is that ULUs do essentially all the work.

This move might seem to obviate any need for an algorithm that could acquire a natural language by hypothesis formation and testing: such an algorithm would be otiose. No significant linguistic generalizations are learned, because none need to be. But in fact Gold's paradox recurs for learning *any* parochial linguistic generalization that involves universal quantification over an unbounded domain, even a lexical generalization. The fact that natural languages are lexically open (see Johnson and Postal, 1980, ch. 14; Postal, 2004; Pullum and Scholz, 2001, 2003) is relevant. Many purely parochial lexical generalizations are highly productive, because it is always possible to add another word to the lexicon of a natural language. Take the English prefix *anti-*, or the suffix *-ish*. It appears that any noun will allow them: we can form words like *anti-borogove*, *anti-chatroom*, *anti-humvee*, *anti-Yushchenko* (though perhaps some of these have not yet been coined); similarly for *borogovish*, *chatroomish*, *humvee-ish*, *Yushchenko-ish*. The Gold problem re-emerges immediately: the indefinitely large set of such derived nouns is a possibly unbounded set of strings, and under Gold's assumptions no algorithm can identify it in the limit from text. For what is the generalization for *anti-*, or for *-ish*? Permissible with all bases beginning with *b*, *ch*, *h*, or *y*? Or all nouns other than *arachnophobe*? Or all non-Latinate roots? Or all bases of two or more syllables? All these and indefinitely many other hypotheses entail the finite corpus of English words considered above. But lexical generalizations are just as underdetermined by the evidence as hypotheses about syntactic structure are, so expanding the evidence won't determine a hypothesis. Yet neither ULUs nor parameters can help here: ex hypothesi these parochial lexical generalizations are just those that are acquired from evidence of use.<sup>10</sup>

Something more than ULUs and various sorts of parameters will be required for the full story about language acquisition. Unless anyone wants to propose the extremely implausible view that no one ever learns anything about any language, we will need a theory of how people learn what *is* learned. And developing such a theory is exactly the non-nativist research program.

If nativists respond to Gold by rejecting learning by hypothesis formation and testing, how do contemporary non-nativists respond? There are many current non-nativist programs, but none of Gold's assumptions are accepted by all of them as relevant to children's first language acquisition:

- Instead of an input of individual strings, the input is taken to be a corpus with rich stochastic information (Lewis and Elman, 2001);
- Instead of acquiring a grammar, the child is assumed to acquire "multiple, simultaneous, probabilistic constraints defined over different types of linguistic and nonlinguistic information" (Seidenberg and MacDonald, 1999, p. 569);

- Instead of only positive data, the child's experience has been investigated and shown to offer plenty of information about what is not in the language (Chouinard and Clark, 2004), and the underdetermination problem is addressed through Bayesian inference, which rules out many accidental generalizations that are supported by the corpus, using probability computations to determine whether certain absences from the corpus are systematic or accidental (Elman, 2005);
- Instead of success being defined as hitting upon a perfectly correct generative grammar, approximative definitions of success have been proposed (the whole field of "probably approximately correct" or "PAC" learning in computer science is based on this move);
- Instead of a hypothesis-testing procedure with whole grammars as hypotheses and only strings as relevant evidence, various monotonic and incremental procedures for approaching a workable grammar are proposed.

The leading challenge these research programs present for linguistic nativism is this: If some of the above proposed methods are utilized by children to learn lexical generalizations, why are ULUs and linguistic parameters required for the acquisition of natural language syntax, but not lexical structure?

## 4 Two Examples of Current Non-nativist Research

So far we have presented some of our reasons for doubting the triumphalist claims made by some linguistic nativists and for doubting that the linguistic nativist/non-nativist dispute is well formed enough to be adjudicated. In what follows, we present two striking research results that have made us optimistic that general cognitive computational mechanisms will explain much more of syntax acquisition than the linguistic nativist has so far acknowledged.

### 4.1 What auxiliary inversion shows

Lewis and Elman (2001) demonstrate that a Simple Recurrent Network (henceforth, SRN) correctly models the acquisition of what linguistic nativists thought required unacquired representations of hierarchical syntactic structure.<sup>11</sup> The case Lewis and Elman consider is the one that Crain (1991) calls the "parade case of an innate constraint." Nativist researchers take it to be one of the strongest arguments for linguistic nativism from the poverty of the stimulus. The reader is typically introduced to the acquisition problem via contrasting pairs of sentences (we cite examples from Laurence and Margolis, 2001, p. 222):

- 4 (a) Ecuador is in South America.
- (b) Is Ecuador in South America?
- 5 (a) The telephone is working again now the storm is over.
- (b) Is the telephone working again now the storm is over?

Given just these four types of sentences, the nativist's assumption is that the child (or is it the linguist?) would be tempted to hypothesize the following kind of syntactic generalization:

- 6 A closed interrogative sentence corresponding to a declarative sentence with *is* as its first verb may be formed by taking the first occurrence of *is* and putting it at the beginning of the sentence, before the subject.

(Set aside for now that (6) contains the linguistic concepts "sentence" and "subject.") The hypothesized (6) turns out to be one of the seductive accidental generalizations that is not supported by further data, as the following pair of sentences shows.

- 7 (a) That woman who is walking her dog is Tom's neighbor.  
 (b) \*Is that woman who walking her dog is Tom's neighbor?

The correct yes/no question formed from (7a) is (8), where the second *is* has been repositioned:

- 8 Is that woman who is walking her dog Tom's neighbor?

The right hypothesis could be framed in various ways, but a straightforward one would be this:<sup>12</sup>

- 9 All closed interrogative clauses formed from declarative clauses, are formed by placing the main clause auxiliary verb at the beginning of the sentence, before the subject.

If this is the child's generalization about the structure of English, then lexical concepts like "main clause" and "auxiliary verb" must, it is supposed, be antecedently known (Fodor, 1981), or the generalization cannot even be entertained. The concept "main clause" relates to hierarchical syntactic structure, not just the linear order of words (the presumed stimulus). So there is every reason for the nativist to suppose that children couldn't frame (9) from stimuli that consist merely of unstructured strings of words.

Certainly, children are reported not to make mistakes like (7b). Crain and Nakayama (1987) ran a study of thirty children (ages 3 to 5 years) who were told to "Ask Jabba if the man who is running is bald." Crain (1991) reports that the outcome was as predicted: children never produced incorrect sentences like (7b). From this Crain (1991) concludes, "The findings of this study then, lend support to one of the central claims of universal grammar, that state that the initial state of the language faculty contains structure-dependence as an inherent property."<sup>13</sup>

If replicable, the finding that 3- to 5-year-old children don't make mistakes like (7b) would be interesting to nativists and non-nativists alike. But it does not support the idea that children have a particular ULU or parameter. And the finding only supports the idea that children have some ULU or other if there is no better explanation.

Could children learn not to make mistakes like (7b) from linguistic stimuli? Chomsky has asserted (without citing evidence) that “A person might go through much or all of his life without ever having been exposed to relevant evidence” of this kind; he states (see Piattelli-Palmarini, 1980, pp. 114–15) that “you can go over a vast amount of data of experience without ever finding such a case” – i.e., a sentence with this structural property. Sampson (1989, 2002) and Pullum and Scholz (2002) question whether such strings are all that rare, providing evidence that relevantly similar strings are found in a variety of texts, including spoken English sources, some of them aimed at fairly young children.<sup>14</sup> But Lewis and Elman (2001) did something particularly interesting that took a different approach.

Lewis and Elman showed that “the stochastic information in the data that is uncontroversially available to children is sufficient to allow learning.” SRNs will “generalize to predict” (2001) in a word-by-word fashion that English has interrogatives like (8), but not like (7b), from training sets that contained strings with the structure of (4) and (5), but not the supposedly rare (8). These training sets encoded “no grammatical information beyond what can be determined from statistical regularities.” Thus, from finite training sets, their SRN does not generalize in hopelessly wrong ways. Nor is learning accomplished by ignoring rich stochastic information in the data.

More recently, Reali and Christiansen (forthcoming) have obtained similar results using a noisy corpus of language addressed to children. These results should be very surprising and intriguing to a linguistic nativist like Crain.

The moral Lewis and Elman draw is that “assumptions about the nature of the input, and the ability of the learner to utilize the information therein, clearly play a critical role in determining which properties of language to attribute to UG” (2001, p. 11). If the relevant stimulus is underestimated to exclude its relevant stochastic features and if the mechanisms of language acquisition are assumed not to exploit them, too much will be taken to be unacquired or triggered. The linguistic nativist seems to tacitly assume that the relevant stimuli for acquisition are simply strings observed in a context of use. But as Lewis and Elman put it: “the statistical structure of language provides for far more sophisticated inferences than those which can be made within a theory that considers only whether or not a particular form appears.” The relevant input includes distributional information about the set of acquisition stimuli (for an SRN, what is in the training set).

Suddenly it begins to look as if what matters for language acquisition is what information is present in the overall stimulus and *how* the stimulus is impoverished, not just *whether* it is impoverished. Lewis and Elman’s training sets included none of the supposedly crucial rare sentences like (8). It begins to seem that structure-dependence can be acquired from the stimuli, even if sentences like (8) are entirely absent in it, contrary to over 25 years of nativist claims.

Of course, there might be other linguistic universals that can’t be learned. But these findings about SRNs raise a series of further questions for both types of research programs. One wants to know exactly which kinds of gaps in training sets SRNs do and do not fill in, and extend this line of work to children’s language acquisition. If children fail to fill in as SRNs do, then that might be grist for the nativist’s mill.

Indeed, the results of these computational experiments suggest that Jerry Fodor’s (1981) claims about the necessity of unacquired linguistic concepts and the impossibility of



learning a language by hypothesis formation and testing only hold for symbolic representations and for the particular learning theory he considers. But they seem irrelevant to the acquisition of distributed representations by means of learning theories based on information-rich statistical regularities in the stimulus, which is a serious contender as a better explanation of the phenomena.<sup>15</sup>

## 4.2 What Simon says

The research program of Newport and Aslin (2000) has found that children might well acquire some morphological/syntactic categories and generalizations from inconsistent and error-ridden data by attending to the stochastic properties of the stimulus. They studied children whose linguistic input is “entirely from speakers who are themselves not fluent or native users of the language” (Newport and Aslin, 2000, p. 12). Their subjects were congenitally and profoundly deaf children acquiring American Sign Language (ASL) as a first language in families with only very imperfect proficiency in ASL. They describe these children’s ASL input as “very reduced and inconsistent.” We will focus on two of their case studies: one involving a child they call Simon and the other involving two children they call Stewart and Sarah.

Simon is widely celebrated in the popular literature on language acquisition; see Pinker (1994, pp. 38–9) for an often cited discussion. Simon’s acquisition of ASL is taken to provide powerful support for linguistic nativism. The case study as reported in Newport and Aslin (2000), however, does not vindicate a nativist interpretation.

Simon is an only child, with deaf parents who were not exposed to ASL until their teens. None of Simon’s teachers knew ASL, so his ASL input was all from his parents. Stewart and Sarah are different in that their parents are hearing, though similar to Simon in other relevant respects. Newport and Aslin report:

Simon’s parents sign are like other late learners: they use virtually all of the obligatory ASL morphemes, but only with middling levels of consistency. On relatively simple morphemes (the movement morphemes of ASL), they average 65–75 percent correct usage. In contrast, Simon uses these morphemes much more consistently (almost 90 percent correct, fully equal to children whose parents are native ASL signers. Thus, when input is quite inconsistent, Simon is nonetheless able to regularize the language and surpass input models. On more difficult morphemes (the hand shape classifiers of ASL), where his parents were extremely inconsistent (about 45 percent correct), Simon did not perform at native levels by age seven; but even there he did surpass his parents. (Newport and Aslin, 2000, p. 13)

Newport and Aslin state competing hypotheses that might explain this finding. The linguistic nativist hypothesis is “that children know, innately, that natural language morphology is deterministic, not probabilistic” (2000, p. 14) and regularize the inconsistent morphological stimuli of their parents signing to conform with ULUs. But they consider one non-nativist alternative: that children have general cognitive mechanisms that “sharpen and regularize” inconsistent patterns in the stimuli.

Newport and Aslin elaborate the latter hypothesis in their discussion of Stewart and Sarah. They note that the correct (native) ASL pattern was used by the parents “with some moderate degree of consistency, while the errors are highly inconsistent”

(2000, p. 19). On the second hypothesis, Simon, Stewart, and Sarah, have acquired ASL from the consistent patterns in their parents overall inconsistent ASL use. This suggests that the overall stochastic information in the inconsistent stimulus is exploited in child language acquisition. That is, learning that is based on the rich stochastic information in the degraded, inconsistent, and messy ASL use of their parents is regularized by children's general stochastic learning mechanisms.<sup>16</sup>

These case studies do not, of course, refute the view stated in (1), that most of the acquisition of natural languages depends on unacquired or unlearned linguistic knowledge. But a clear moral is that without careful attention to the stochastic properties of the stimulus, the hypothesis that general cognitive mechanisms play a significant role in language acquisition is *not* ruled out. Perhaps because of the way Newport and Aslin's research has been publicized, the finding that Simon regularized over inconsistent input has been taken as clear support for linguistic nativism by means of a poverty of the stimulus argument. But this interpretation is premature. It looks as if the relative frequencies of patterns and inconsistencies in the overall stimulus is more important than the mere fact that it contains errors or is inconsistent.<sup>17</sup> Children have unacquired and unlearned mechanisms that regularize partial patterns that they detect – whether linguistic or not. That (if true) is certainly an inherent fact about children that assists in their language acquisition; but it does not imply possession of ULUs.

## 5 Conclusion

We have argued that, as of today, to maintain that linguistic nativism has triumphed over non-nativism demands tacitly accepting at least one of three rather unappealing views.

The first is to accept that linguistic nativism is the negation of reductive empiricism – that is, to depict all contemporary non-nativists as defenders of John Locke and B. F. Skinner – and declare victory. But that will necessitate ignoring the content of actual contemporary non-nativist theories.

The second is to take it on faith that one day an appropriate sense of “innate” will be discovered that makes it a coherent and contentful theoretical term with all the relevant specialized language acquisition mechanisms in its extension. But the meanings of “innate” that are in current use in linguistics are not all empirically equivalent, and the currently hypothesized mechanisms of language acquisition do not fall under a definite concept.

And the third strategy is to posit parameters, set by triggering (in some sense of “triggering”), for absolutely every parochial peculiarity of every natural language, even lexical generalizations. But if the set of posited parameters tracks the set of parochial features of natural languages, the theory is rendered vacuous as a theory of language acquisition: instead of an explanation of how language is acquired we get just a list of ways natural languages can differ.

None of these three strategies looks productive to us. But the defender of the claim that linguistic nativism has vanquished rival non-nativist views is in the unfortunate position of tacitly accepting at least one of them.

## Acknowledgments

This work was supported in part by the Radcliffe Institute for Advanced Study at Harvard University, where the authors completed the final version of the chapter. We are grateful to Eve Clark, Shalom Lappin, Ivan Sag, and Stuart Shieber for useful discussions and correspondence; to Claire Creffield for highly intelligent copy-editing assistance; and to Jeff Elman, Mark Liberman, Diana Raffman, J. D. Trout, and Robert Stainton for their valuable comments on an earlier draft. No one mentioned here should be assumed to agree with the views we express, of course.

## Notes

- 1 Before you charge us with being unfair, take a look at some quotes: “Chomsky’s demonstration . . . is *the* existence proof for the possibility of a cognitive science” (Fodor, 1981, p. 258); “How can a system of ideas like this have succeeded in capturing the intellectual allegiance of so many educated people?” (Sampson, 1997, p. 159); “A glance at any textbook shows that . . . generative syntax has uncovered innumerable such examples” (Smith, 1999, p. 42); “Is the idea supposed to be that there is no (relevant) difference between my granddaughter, her pet kitten, a rock, a chimpanzee?” (Chomsky, quoted by Smith, 1999, pp. 169–70); “Her rhetorical stance . . . invites comparison with creationists’ attacks on the hegemony of evolution” (Antony, 2001, p. 194, referring to Cowie, 1999). There is more wild-eyed stuff where this came from, and it is spouted by both sides.
- 2 Several commentators seem to have assumed this about Pullum and Scholz 2002, which is actually framed as an effort at stimulating nativists to present evidence that would actually count in favor of their view. Scholz and Pullum 2002 replies to several critics and tries to make the goals clearer.
- 3 Of course, it’s premature to celebrate a non-nativist victory too. Geoffrey Sampson’s announcement in a newspaper article that nativism has collapsed (Sampson, 1999) is an example of excess at the other extreme.
- 4 We note that Hauser, Chomsky, and Fitch, 2002, claim that the core language faculty comprises just “recursion” and nothing else, apparently accepting such a pyrrhic nativism; but they are answered on this point in great detail by Pinker and Jackendoff 2005.
- 5 “The linguistic literature is full of arguments of this type” (Lightfoot, 1998, p. 585); “A glance at any textbook shows that half a century of research in generative syntax has uncovered innumerable such examples” (Smith, 1999, p. 42).
- 6 Notice, we take “linguistic nativism” to denote a claim, not just a strategy. Some psycholinguists clearly differ. Fodor and Crowther 2002, for example, thinks linguistic nativism is a methodology that “assumes everything to be innate that could be innate.” This would presumably contrast with a non-nativist methodology that assumes everything to be acquired that could be acquired. But these are not the forms of linguistic nativism and non-nativism we address.
- 7 The principles and parameters approach is basically abandoned in the controversial recent development known as the “minimalist program”; see Pinker and Jackendoff 2005 for a critique from a standpoint that is decidedly skeptical but nonetheless clearly nativist.
- 8 “Parochial” here means varying between natural languages, rather than being true of all of them.
- 9 “Interesting” is used here in a sense stemming from formal language theory, where finite or specially gerrymandered classes are not interesting, but broad and mathematically natural classes such as the regular languages or the context-free languages are interesting.

- An excellent introduction to both the mathematics and the linguistic and psycholinguistic relevance can be found in Levelt 1974.
- 10 Janet Fodor 1989 wrestles with this issue, without arriving at a satisfying resolution. See Culicover (1999, p. 15) for remarks with which we agree: “Since human beings acquire both the general and the idiosyncratic, there must be a mechanism or mechanisms that can accommodate both . . . Even if we assume that the most general correspondences are instantiations of linguistic universals that permit only simple parametric variation, the question of how the rest of linguistic knowledge is acquired is left completely unexplored.”
  - 11 The SRN is a “three-layer feed-forward network – made up of the input, hidden, and output layers – augmented by a context layer.” We should note that Prinz 2002 – a work which unfortunately we encountered only after this chapter was almost completed – describes Elman’s work as showing “that a dumb pattern detector can pick up on structural relations” (p. 206). This seems overstated. Prinz seems unaware of the growing literature suggesting an alternative interpretation: that children are very capable and sophisticated learners of transition probabilities. Elman’s computational models are particularly important in light of the discovery of children’s stochastic learning capacities.
  - 12 We’re ignoring one complication, as other discussions generally do: if there is no main clause auxiliary verb, the auxiliary verb *do* is required.
  - 13 There is actually a great deal to be made clear about just what the higher-order property of “structure-dependence” is. The statement in (9) is not universal: other languages do not form interrogative sentences in the same way as English. What could perhaps be universal is some metaprinciple about the form of suitable candidates for tentative consideration as principles of grammar. No one has ever really made this precise. We will ignore the matter here.
  - 14 It is worth pointing out that there is a deep inconsistency in the nativist literature concerning the kind of stimulus that is relevant to showing that the stimulus for auxiliary inversion is impoverished. On the one hand, nativists often claim that only the characteristics of child directed speech are relevant for child language acquisition, since children acquire language primarily from child-directed speech. On the other hand, it is often pointed out that in some cultures adults do not direct speech to children until they are verbally fluent, so *ex hypothesi*, in these cultures the relevant stimulus is not speech directed specifically toward children. The reason this is important is that how impoverished the stimulus is depends on what stimuli are counted as relevant. For an informed discussion see Clark 2003.
  - 15 Fodor 1981 ignores stochastically based learning. Pessimistic results like those of Gold 1967 simply do not apply under the assumption that linguistic input is modeled as a stochastic process and not text (see Scholz, 2004). Elsewhere, Fodor claims that stochastic learning can do nothing but recapitulate the distributional properties of the input. Elman’s SRN is a counterexample to that claim. However, this is not the place to reply to Fodor’s criticisms of connectionism.
  - 16 Prinz (2002, pp. 209–10) takes Singleton and Newport (2004) to show that “children can turn a moderate statistical regularity into a consistently applied rule.” But Simon’s regularization of morphological structure need not be seen as evidence for the acquisition of rules in the linguist’s sense. Simon’s accomplishment is to exhibit more statistical regularity than his input did. This does not of itself tell us that any rules were acquired (though they might have been).

There is more to be said about sign language, though we have insufficient space here. Findings about language acquisition from inconsistent stimuli have played an important role in research on cross-cohort syntactic change in Nicaraguan Sign Language where there is iterated regularization across successive stages in the development of the

language. Here part of what is being investigated is syntactic regularization across a range of different inconsistent stimuli. And this line of research promises to provide insights into creolization, another controversial topic relating to language acquisition (see Bickerton, 1984).

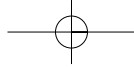
- 17 Notice, inconsistency and error were *not* features of the input to the learner considered by Lewis and Elman; they assumed that the consistent and correct data lacked instances of one particular kind of sentence.

## References

- Antony, L. M. (2001). Empty heads. *Mind and Language*, 16, 193–214.
- Ariew, A. (1996). Innateness and canalization. *Philosophy of Science Supplement*, 63, 19–27.
- (1999). Innateness is canalization: A defense of a developmental account of innateness. In V. Hardcastle (ed.), *Biology Meets Psychology: Conjectures, Connections, Constraints*. Cambridge, MA: MIT Press.
- (2003). Innateness and triggering: Biologically grounded nativism. Manuscript, University of Rhode Island.
- Aslin, R. N., Saffran, J. R., and Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9, 321–4.
- Barkow, J., Cosmides, L., and Tooby, J. (1992). *The Adapted Mind*. Oxford: Oxford University Press.
- Bateson, P. P. G. (ed.) (1991) *The Development and Integration of Behaviour*. Cambridge: Cambridge University Press.
- (2004) The origins of human differences. *Daedalus*, 133/4, 36–46.
- Bealer, G. (1999). The a priori. In J. Greco and E. Sosa (eds.), *The Blackwell Guide to Epistemology*. Oxford: Blackwell.
- Bickerton, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences*, 7, 173–221.
- Block, N. (1981). *Readings in the Philosophy of Psychology*, vol. 2. Cambridge, MA: Harvard University Press.
- Boyd, R. N. (1981). Scientific realism and naturalistic epistemology. In P. Asquith and R. Giere (eds.), *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1980*, vol. 2, 613–62. East Lansing, MI: Philosophy of Science Association.
- (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61, 127–48.
- Chomsky, N. (2000). *The Architecture of Language*. New Delhi: Oxford University Press.
- Chouinard, M. and Clark, E. (2004). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30, 637–69.
- Christiansen, M. H. and Curtin, S. L. (1999). The power of statistical learning: No need for algebraic rules. *Proceedings of the 21st Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.
- Clark, E. (2003). *First Language Acquisition*. Cambridge: Cambridge University Press.
- Cowie, F. (1999). *What's Within? Nativism Reconsidered*. New York: Oxford University Press.
- Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, 14, 597–650.
- Crain, S. and Nakayama, M. (1987). Structure dependence in grammar formation. *Language*, 63, 522–43.

- Culicover, P. W. (1999). *Syntactic Nuts: Hard Cases, Syntactic Theory, and Language Acquisition*. Oxford: Oxford University Press.
- Elman, J. L. (2003). Generalization from sparse input. *Proceedings of the 38th Annual Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society.
- (2005). Connectionist models of cognitive development: Where next? *Trends in Cognitive Sciences*, 9, 111–17.
- Fodor, J. A. (1981). The current status of the innateness controversy. In *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press.
- Fodor, J. D. (1989). Learning the periphery. In R. J. Matthews and W. Demopoulos (eds.), *Learnability and Linguistic Theory*. Dordrecht: Kluwer.
- Fodor, J. D. and Crowther, C. (2002). Understanding stimulus poverty arguments. *The Linguistic Review*, 19, 105–45.
- Gibson, E. and Wexler, K. (1994). Triggers. *Linguistic Inquiry*, 25, 407–54.
- Gold, E. M. (1967). Language identification in the limit. *Information and Control*, 10, 447–74.
- Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431–6.
- Griffiths, P. E. (2002). What is innateness? *The Monist*, 85, 70–85.
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–79.
- Jain, S., Osherson, D., Royer, J. S., and Sharma, A. (1999). *Systems That Learn* (2nd edn.). Cambridge, MA: MIT Press.
- Johnson, D. E. and Postal, P. M. (1980). *Arc Pair Grammar*. Princeton, NJ: Princeton University Press.
- Kornblith, H. (1993). *Inductive Inference and Its Natural Ground: An Essay in Naturalistic Epistemology*. Cambridge, MA: MIT Press.
- Laurence, S. and Margolis, E. (2001). The poverty of the stimulus argument. *British Journal for the Philosophy of Science*, 52, 217–76.
- Levelt, W. J. M. (1974). *Formal Grammars in Linguistics and Psycholinguistics* (3 vols). The Hague: Mouton.
- Lewis, J. D. and Elman, J. L. (2001). Learnability and the statistical structure of language: Poverty of stimulus arguments revisited. *Proceedings of the 26th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Li, M. and Vitányi, P. (1997). *An Introduction to Kolmogorov Complexity and Its Applications* (2nd edn.). New York: Springer.
- Lidz, J., Gleitman, H., and Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of Universal Grammar. *Cognition*, 87, 151–78.
- Lightfoot, D. (1998). Promises, promises: General learning algorithms. *Mind and Language*, 13, 582–7.
- Lipton, P. (1991). *Inference to the Best Explanation*. New York: Routledge.
- (1998). Induction. In M. Curd and J. A. Cover (eds.), *Philosophy of Science: The Central Issues*. New York: Norton. (Reprint of a passage from Lipton, 1991.)
- Marler, P. (1999). On innateness: Are sparrow songs “learned” or “innate”? In M. D. Hauser and M. Konishi (eds.), *The Design of Animal Communication*. Cambridge, MA: MIT Press.
- Matthews, R. J. and Demopoulos, W. (eds.) (1989). *Learnability and Linguistic Theory*. Dordrecht: Kluwer.
- Newport, E. L. and Aslin, R. N. (2000). Innately constrained learning: blending old and new approaches to language acquisition. In S. C. Howell, S. A. Fish, and T. Keith-Lucas (eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

- (2003). Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48, 127–62.
- Piattelli-Palmarini, M. (1980). *Language and Learning: The Debate Between Jean Piaget and Noam Chomsky*. London: Routledge and Kegan Paul.
- Pinker, S. (1994). *The Language Instinct: How the Mind Creates Language*. New York: William Morrow.
- Pinker, S. and Jackendoff, R. S. (2005). The faculty of language: What's special about it? *Cognition*, 95, 201–36.
- Postal, P. M. (2004). The openness of natural languages. In P. M. Postal, *Skeptical Linguistic Essays*. New York: Oxford University Press.
- Prinz, J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: MIT Press.
- Pullum, G. K. and Scholz, B. C. (2001). On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In P. de Groot, G. Morrill, and C. Retoré (eds.), *Logical Aspects of Computational Linguistics: 4th International Conference*. Berlin: Springer Verlag.
- (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, 19, 9–50.
- (2003). Linguistic models. In M. T. Banich and M. Mack (eds.), *Mind, Brain, and Language: Multidisciplinary Perspectives*. Mahwah, NJ: Erlbaum.
- Putnam, H. (1963a). Degree of confirmation and inductive logic. In P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*. La Salle, IL: Open Court. (Reprinted in Putnam, 1979).
- (1963b). Probability and confirmation. *The Voice of America, Forum Lectures, Philosophy of Science Series*, 10. Washington DC: US Information Agency. (Reprinted in Putnam, 1979).
- (1979). *Mathematics, Matter and Method: Philosophical Papers*, vol. 1 (2nd edn.). New York: Cambridge University Press.
- Reali, F. and Christiansen, M. (forthcoming). Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. *Cognitive Science*.
- Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–8.
- Saffran, J. R. and Wilson, D. P. (2003). From syllables to syntax: Multi-level statistical learning by 12-month-old infants. *Infancy*, 4, 273–84.
- Sampson, G. (1989). Language acquisition: Growth or learning? *Philosophical Papers*, 18, 203–40.
- (1997). *Educating Eve*. London: Cassell.
- (1999). Collapse of the language nativists. *The Independent* (April 9), 7.
- (2002). Exploring the richness of the stimulus. *The Linguistic Review*, 19, 73–104.
- Samuels, R. (2002). Nativism in cognitive science. *Mind and Language*, 17, 233–65.
- Scholz, B. C. (2004). Gold's theorems and the logical problem of language acquisition. *Journal of Child Language*, 31, 959–61.
- Scholz, B. C. and Pullum, G. K. (2002). Searching for arguments to support linguistic nativism. *The Linguistic Review*, 19, 185–223.
- Seidenberg, M. S. and MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, 23, 569–88.
- Singleton, J. L. and Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology*, 49, 370–407.
- Smith, N. (1999). *Chomsky: Ideas and Ideals*. Cambridge: Cambridge University Press.



- Sober, E. (1998). Innate knowledge. In E. Craig (ed.), *Routledge Encyclopedia of Philosophy*, vol. 4. London: Routledge.
- Stich, S. (1975). *Innate Ideas*. Berkeley, LA: University of California Press.
- Trout, J. D. (1998). Nativism, statistical complexity, and speech. Presented at the 24th Annual Meeting of the Society for Philosophy and Psychology, Minneapolis, MN.
- Waddington, C. H. (1940). *Organizers and Genes*. Cambridge: Cambridge University Press.

