

Irrational Nativist Exuberance*

Barbara C. Scholz

San José State University

Geoffrey K. Pullum

University of California, Santa Cruz

Draft of April 11, 2005. To appear in Rob Stainton (ed.), *Debates in Cognitive Science* (Oxford: Basil Blackwell, in press). Please do not quote in print without contacting the authors (pullum@ucsc.edu, bscholz@email.sjsu.edu).

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.¹ In this chapter, however, we draw a firm line 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 involved here.

We are skeptical about recent triumphalist claims for linguistic nativism, and this may lead to us being mistaken for defenders of empiricism.² 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),³ 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 the syntax, the reader will note that we avoid using ‘innate’ whenever we can. 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. Whatever that involves, we assume without argument that under linguistic nativism these constraints are (i) either universal linguistic principles or constraining biases introduced by language-specialized cognitive mechanisms, and (ii) either unacquired or acquired but not learned. The non-nativist, by contrast, holds that the constraints on language acquisition do not meet conditions (i) and (ii).

However, 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. Nativist/non-nativist disputes about syntax acquisition concern *what* constrains or biases language acquisition, not whether it is constrained or biased.

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, so it is the view that at least one thing is in the intellect that was *not* acquired from the senses. This is too weak to be an interesting form of contemporary nativism. Contemporary linguistic nativism claims that more than just one language specific unacquired cognitive mechanism or contentful principle is required for language acquisition.⁴ 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. It would also be a mystery why nativist linguists have attempted to establish the existence of so many such principles.⁵

For the purposes of this chapter, linguistic nativism is 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 alone. The research program of linguistic nativism aims to show, proposition by proposition (and mechanism by mechanism) that little knowledge of syntactic structure is acquired or learned from sensory stimuli by means of general cognitive mechanisms and non-cognitive mechanisms (whether specialized for language or not). Thus the discovery of one (or even a few) special 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 rejects (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 (section 2), and how unacquired language-specialized mechanisms solve certain problems of language acquisition (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 during the Enlightenment and B. F. Skinner in the first half of the 20th century. 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 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. (1981:273)

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 concept acquisition is 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: 267).

Given this reductive characterization, Fodor struggles to specify the point of disagreement between nativists and both historical empiricists and contemporary non-nativists (1981:279-283). 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 of such areas as concept acquisition 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) appears to be an application of this idea to the psychological task of 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 acquisition mechanisms (see Block 1981: 280). But Chomsky (2000: 50) presses this point to a parodic extreme when he writes:

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.

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 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 representationalism, 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. (forthcoming) 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*.” Elman (2003) describes this error as taking 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. It is 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 what kinds of stochastic patterns infants can acquire (Saffran, Aslin and Newport 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 (Elman 2003). Elman, Newport, and others have shown that the relevant features of the stimulus for statistically-based language acquisition models are the stochastic properties of the overall input. These relevant properties of the overall input are not exhausted by the presence or absence of individual strings therein.

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). Non-nativists can whole-heartedly accept that non-cognitive language-specialized mechanisms and general cognitive computational mechanisms constrain what is acquired; language acquisition is, for example, unquestionably constrained by quantitative resource bounds and sensory channel limits.

2. What innateness is

The hypothesis that some features of natural languages are acquired by triggering is characteristic of the ‘principles and parameters’ theory.⁶ Parameters are supposed to “reduce the difficulty of the learning problem” (Gibson and Wexler: 1994: 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.⁷ 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 both unacquired and unlearned. 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 innate.

On one side of the debate, Samuels (2002), developing J. A. Fodor’s view, 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 2002:246). 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 (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, has argued that innateness is **environmental canalization**: on which “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 vs. triggered vs. 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 thirty 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: 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 1975:15), the parameter is not merely set off or activated. 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: 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: 409–410): 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 trigger 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 1994:408, n.2).

Instantaneous acquisition is a brute-causal psychological process that is 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 Ariew’s sense.

Doubtless Gibson and Wexler would reject both Samuels' and Ariew'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 help Gibson and Wexler: 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 language.

Some scientists talk as if what is **universal** across all typical members of the species or all across all natural languages is innate (see, e.g., Barkow, Cosmides and Tooby 1992); but neither the products of instantaneous acquisition nor those of accelerated acquisition are universal.

Stich (1975:12) suggests a **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 not be tempted to revert to the idea that what is innate is known *a priori*. First, if *a priori* knowledge is defined as "what is known independently of 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. Second, if the products of accelerated acquisition are claimed to be innate because they are known *a priori*, then it would appear that we are being given an epistemological concept of innateness when what was wanted was a psychological one.

Bealer (1999) has more recently approached the concept of the *a priori* through *a priori* evidence, "not imparted through experiences but rather through intuitions" (245). According to Bealer, "For you to have an intuition that A is just for it to *seem* to you that A" (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 products of accelerated acquisition are arguably imparted through experiences. Bealer's idea might be pursued in linguistics if very young children and infants have conscious cognitive intuitions that reliably track the products of accelerated acquisition, since cognitive intuitions, not perceptual experiences or sensory intuitions are *a priori* evidence on Bealer's account. It is an empirical question whether children have this kind of intuitions, and part of what would need to be shown is that infants are not tracking the distributional frequencies of sensory stimuli.

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 play a role in the explanation of language acquisition. Rather, our worry is that treating these distinct empirically distinct 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: 37–39) 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, and it is worth pointing out that it is not just non-nativists who propose jettisoning the term ‘innate’. Marler (1999) advocates abandoning the use of both ‘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. (1999: 311)

Marler is not advocating some bland form of interactionism. Rather, he is suggesting that the innate-learned dichotomy functions to ‘black-box’ relevant developmental factors in way that undermines our understanding of development.

If our recommendation is followed, then the partisan nativist/non-nativist dispute will, we think, be less polarized, and the partisan debates will focus on which unacquired language-specialized cognitive mechanisms play a role in language acquisition. In the following section, we critically examine the alleged payoff for language acquisition of hypothesizing such unacquired linguistic universals.

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 (2002: 1577) 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.

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” (see Curd and Cover 1998: 413), stressing the necessity of ULUs; and Laurence and Margolis (2001:221) stress that non-nativists cannot countenance biases of any sort: “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 ULUs. 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 inductive 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 Gold (1967).

The idea of studying inductive inference by investigating the limits of an abstract pattern-learning machine seems to originate with Putnam (1963a, 1963b). Putnam further proved a basic impossibility result about what such a machine could accomplish (namely, that there could be no “perfect” learning machine that would learn any presented regularity), and laid out the strategy that would later be used to prove the main result of Gold (1967). However, it was Gold’s work that stimulated the rise to prominence 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, Gold advocated investigation of “the limiting behavior of the guesses as successively larger and larger bodies of information are considered” (Gold 1967: 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 grammars⁸ the learning problem unsolvable, because what has to be learned is deductively underdetermined by the positive evidence (evidence about what *is* in the language; learning from such evidence is called “identification in the limit from text”). On the other hand, he showed that if the evidence includes both *what is and what is not* in the language, the learning problem is solvable.

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, the result is 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 for 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], [tʃ], [h], or [j]? Or all nouns other than *arachnophobe*? Or all non-Latinate roots? Or all *n*-syllable roots for $n \leq 3$? All these and indefinitely many other hypotheses are supported by the foregoing examples. Lexical generalizations are just as underdetermined by the evidence as hypotheses about syntactic structure are. Yet neither ULUs nor parameters can help here: *ex hypothesi* these parochial lexical generalizations are just those that are acquired from evidence of use.⁹

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 statistical structure. 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). And 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 that if some of the above proposed methods are utilized by children to learn lexical generalizations, why are unacquired syntactic generalizations required for the acquisition of natural language syntax?

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 language 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.¹⁰ 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 2001:222):

- (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 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 statement in (6) turns out to be one of the seductive accidental generalizations that if hypothesized by the language learner 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:¹¹

- (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 acquire linguistic knowledge of hierarchical structure from stimuli that consists merely of unstructured strings of words.

Certainly, children don't 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, "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."¹²

The fact that 3 to 5 year-old children don't make mistakes like (7b) should be interesting to nativists and non-nativists alike. But it does not show that any particular linguistic principle is unacquired, or parameter activated.

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: 114–115) 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.¹³ 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" (p. 8) in a word-by-word fashion that English has interrogatives like (8), but not like (7b), from training sets that contained strings like those in (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. This result should be very surprising and intriguing to a linguistic nativist like Crain, since the exact character of the SRN's architectural constraints are critical.

The moral Lewis and Elman draw from their research 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" (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. The 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 both what the stimulus is 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 rare, contrary to over twenty-five years of nativist claims.

Of course, there might be other linguistic universals that can’t be learned. But this finding about SRNs raises 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. But note that Lewis and Elman did not show that SRNs formulate hypotheses like (9) that require the antecedent acquisition of concepts like ‘main clause’ and ‘sentence’ that they then test.

Indeed, their work suggests 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.¹⁴

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” (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: 38–39) for a typical 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 the only child of deaf parents who were not exposed to ASL until their teens. None of Simon’s teachers know ASL, so his ASL input is 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 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% correct usage. In

contrast, Simon uses these morphemes much more consistently (almost 90% 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% correct), Simon did not perform at native levels by age 7; but even there he did surpass his parents. (2000: 13)

Newport and Aslin state competing hypotheses that might explain this finding. One is nativist and one is not. The nativist hypothesis is “that children know, innately, that natural language morphology is deterministic, not probabilistic” (p. 14) and regularize the inconsistent morphological stimuli of their parents signing to conform with this unacquired contentful information. The other is that children have general cognitive mechanisms that “sharpen and regularize” inconsistent patterns the stimuli.

Newport and Aslin elaborate the latter hypothesis in their discussion of Stewart and Sara. 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” (p. 19). On the second hypothesis, Simon, Stewart, and Sara, have acquired ASL from the consistent patterns in their parents overall inconsistent ASL use. This suggests that the overall distributional frequencies 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.

These case studies do not, of course, refute the nativist hypothesis that children have unacquired knowledge that morphology is deterministically regular or the hypothesis that they have parameters that are triggered by the stimulus. But a clear moral is that without careful attention to the stochastic properties of the stimulus, learning 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 nativism by means of a poverty of the stimulus argument. But this interpretation is premature. It begins to look like for language learning the frequencies of degraded patterns in the overall stimulus are more important than the fact that it contains errors or is inconsistent.¹⁵

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 sequences of data that is inconsistent overall. Here what is being investigated is syntactic regularization across a range of different inconsistent input circumstances. And this line of research promises to provide insights into creolization, another topic of dispute between partisan nativists and non-nativists (Bickerton 1984).

Conclusion

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

The first is to define nativism as 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 research, since it doesn’t fit this stereotype at all.

The second is to take it on faith that one day an appropriate sense will be given to 'innate' that will make it a coherent 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 they cross-cut the currently hypothesized mechanisms of language acquisition. If the term does not correspond to a unique concept, there will be no definite content to the empirical generalizations that employ it.

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 having to adopt at least one of them.

Notes

* We are very grateful to Eve Clark, Jeff Elman, Mark Liberman, Diana Raffman, Ivan Sag, J. D. Trout, and Robert Stainton for their comments on an earlier draft.

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 1981:258); “How can a system of ideas like this have succeeded in capturing the intellectual allegiance of so many educated people?” (Sampson 1997:159); “A glance at any textbook shows that . . . generative syntax has uncovered innumerable such examples” (Smith 1999:42); “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 1999:169–170); “Her rhetorical stance . . . invites comparison with creationists’ attacks on the hegemony of evolution” (Antony 2001: 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 that Pullum and Scholz (2002) is an argument for empiricism, when it 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 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 Notice, we take ‘linguistic nativism’ to denote a claim, not just a strategy. Some psycholinguists clearly differ. Janet Fodor (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.

⁶ 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.

⁷ ‘Parochial’ here means capable of varying between natural languages, rather than being true of all of them.

⁸ ‘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).

⁹ Janet Fodor (1989) wrestles with this issue, without arriving at a satisfying resolution. See Culicover (1999:15) 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.”

¹⁰ The SRN is a “three-layer feed-forward network — made up of the input, hidden, and output layers — augmented by a context layer.”

¹¹ We’re ignoring one complication, as other discussions generally do: if there is no main clause auxiliary verb, the auxiliary verb *do* is required.

¹² 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.

¹³ 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 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).

¹⁴ 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.

¹⁵ 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, Louise M. (2001) Empty heads. *Mind & Language* **16**: 193–214.
- Ariew, André (1996) Innateness and canalization. *Philosophy of Science Supplement* 63:19-27.
- Ariew, André (1999) Innateness is canalization: a defense of a developmental account of innateness. In Valerie Hardcastle (ed.), *Biology Meets Psychology: Conjectures, Connections, Constraints*, 117–138. Cambridge, MA: MIT Press.
- Ariew, André (2003) Innateness and triggering: Biologically grounded nativism. MS, University of Rhode Island.
- Barkow, J., Cosmides, L., and Tooby, J. (1992) *The Adapted Mind*. Oxford: Oxford University Press.
- Bealer, George (1999) The a priori . In John Greco and Ernest Sosa (eds.), *The Blackwell Guide to Epistemology*. Oxford: Basil Blackwell.
- Bickerton, Derek (1984) The language bioprogram hypothesis. *Behavioral and Brain Sciences* **7**: 173–221.
- Block, Ned (1981) *Readings in the Philosophy of Psychology, Vol. 2*. Cambridge, MA: Harvard University Press.
- Boyd, Richard N. (1981) Scientific realism and naturalistic epistemology. In Peter Asquith and Ronald Giere (eds.), *PSA 1980, Vol. 2*: 613–662. East Lansing, MI: Philosophy of Science Association.
- Boyd, Richard N. (1991) Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies* **61**:127–148.
- Chomsky, Noam (2000) *The Architecture of Language*. New Delhi: Oxford University Press.
- Chouinard, Michelle, and Clark, Eve. (2004) Adult reformulations of child errors as negative evidence. *Journal of Child Language*. **30**: 637-669.
- Christiansen, M.H. & Curtin, S.L. (1999). The power of statistical learning: No need for algebraic rules. In The Proceedings of the 21st Annual Conference of the Cognitive Science Society, 114-119. Mahwah, NJ: Lawrence Erlbaum Associates.
- Clark, Eve (2003) *First Language Acquisition*. Cambridge: Cambridge University Press.
- Cowie, Fiona (1999) *What's Within? Nativism Reconsidered*. New York: Oxford University Press.
- Crain, Stephen (1991) Language acquisition in the absence of experience. *Behavioral and Brain Sciences* **14**: 597–650.
- Crain, Stephen and Mineharu Nakayama (1987) Structure dependence in grammar formation. *Language* **63**: 522–543.
- Culicover, Peter W. (1999) *Syntactic Nuts: Hard Cases, Syntactic Theory, and Language Acquisition*. Oxford: Oxford University Press.
- Elman, Jeffrey L. (2003) Generalization from sparse input. *Proceedings of the 38th Annual Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society.
- Elman, Jeffrey L. (2005) Connectionist models of cognitive development: where next? *Trends in Cognitive Sciences* **9**: 111–117.
- Fodor, Janet Dean (1989) Learning the periphery. In Matthews and Demopoulos (eds.), 129–154.
- Fodor, Jerry A. (1981) The current status of the innateness controversy. In *Representations*, 257–316. Cambridge, MA: MIT Press.
- Gibson, Edward, Wexler, Kenneth (1994) Triggers. *Linguistic Inquiry* **25** 407-454.
- Gold, E. Mark (1967) Language identification in the limit. *Information and Control* **10**: 447–474.

- Griffiths, Paul 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, Sanjay; Daniel Osherson; James S. Royer; and Arun Sharma (1999) *Systems That Learn*, second edition. Cambridge, MA: MIT Press.
- Johnson, David E. and Paul M. Postal (1980) *Arc Pair Grammar*. Princeton, NJ: Princeton University Press.
- Kornblith (1993) *Inductive Inference and Its Natural Ground: An Essay in Naturalistic Epistemology*. Cambridge, MA: MIT Press.
- Laurence, Stephen and Eric Margolis (2001) The poverty of the stimulus argument. *British Journal for the Philosophy of Science* **52**: 217–276.
- Levelt, W. J. M. (1974) *Formal Grammars in Linguistics and Psycholinguistics* (3 volumes). The Hague: Mouton.
- Lewis, John D. and Jeffrey L. Elman (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.
- Lidz, Jeffrey, Henry Gleitman, and Lila Gleitman (2003). Understanding How Input Matters: Verb Learning and the Foot-Print of Universal Grammar. *Cognition* **87**: 151-178.
- Lipton, Peter (1991) *Inference to the Best Explanation*. New York: Routledge. [Relevant passage reprinted as ‘Induction’ in Martin Curd and J. A. Cover (eds.), *Philosophy of Science: The Central Issues* (New York: W. W. Norton, 1998), 412–425.]
- Marler, Peter (1999) On Innateness: Are Sparrow Songs “Learned” or “Innate”? In Marc D. Hauser and Mark Konishi (eds), *The Design of Animal Communication*, Cambridge, MA: MIT Press.
- Matthews, Robert J. and William Demopoulos, eds. (1989) *Learnability and Linguistic Theory*. Dordrecht, Netherlands: Kluwer Academic.
- Newport, Elissa L. and Richard N. Aslin (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*, 1–21. Somerville, MA: Cascadilla Press.
- Newport, Elissa L. and Richard N. Aslin (2003) Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology* **48**: 127–162.
- Piattelli-Palmarini, Massimo (1980) *Language and Learning: The Debate Between Jean Piaget and Noam Chomsky*. London: Routledge and Kegan Paul.
- Pinker, Steven (1994) *The Language Instinct: How the Mind Creates Language*. New York: William Morrow.
- Pinker, Steven and Ray S. Jackendoff (2005) ‘The faculty of language: what’s special about it?’ *Cognition* **95**: 201–236.
- Postal, Paul M. (2004) ‘The openness of natural languages.’ Chapter 6 of *Skeptical Linguistic Essays*, 173–201. New York: Oxford University Press.
- Pullum, Geoffrey K. and Barbara C. Scholz (2001) On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In *Logical Aspects of Computational Linguistics: 4th International Conference* (Lecture Notes in Artificial Intelligence, 2009), ed. by Philippe de Groote, Glyn Morrill, and Christian Retoré, 17–43. Berlin: Springer Verlag.
- Pullum, Geoffrey K. and Barbara C. Scholz (2002) Empirical assessment of stimulus poverty arguments. *The Linguistic Review* **19**: 9–50.

- Pullum, Geoffrey K. and Barbara C. Scholz (2003) Linguistic models. In *Mind, Brain, and Language: Multidisciplinary Perspectives*, ed. by Marie T. Banich and Molly Mack, 113–141. Mahwah, NJ: Lawrence Erlbaum.
- Putnam, Hilary (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), 270–292.
- Putnam, Hilary (1963b) ‘Probability and confirmation.’ *The Voice of America, Forum Philosophy of Science*, 10. Washington DC: U.S. Information Agency. Reprinted in Putnam (1979), 293–304.
- Putnam, Hilary (1979) *Mathematics, Matter and Method: Philosophical Papers, Volume 1*, second edition. New York: Cambridge University Press.
- Saffran, Jenny R., Richard N. Aslin and Elissa L. Newport (1996) Statistical learning by 8-month-old infants. *Science* **274**: 1926–1928.
- Sampson, Geoffrey (1989) Language acquisition: growth or learning? *Philosophical Papers* **18**: 203–240.
- Sampson, Geoffrey (1997) *Educating Eve*. London: Cassell.
- Sampson, Geoffrey (1999) Collapse of the language nativists. *The Independent* (April 9), 7.
- Sampson, Geoffrey (2002) Exploring the richness of the stimulus. *The Linguistic Review* **19**, 73–104.
- Samuels, Richard (2002) Nativism in cognitive science. *Mind & Language* **17**: 233–265.
- Scholz, Barbara (2004) Gold’s theorems and the logical problem of language acquisition. *Journal of Child Language* **31**: 959–961.
- Scholz, Barbara C. and Geoffrey K. Pullum (2002) Searching for arguments to support linguistic nativism. *The Linguistic Review* **19**: 185–223.
- Smith, Neil (1999) *Chomsky: Ideas and Ideals*. Cambridge University Press.
- Sober, Elliott (1998) Innate knowledge. In Edward Craig (ed.), *Routledge Encyclopedia of Philosophy, Vol. 4*, 794–797.
- Stich, Stephen (1975) *Innate Ideas*. Berkeley 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.