# Stanford Encyclopedia of Philosophy

# Innateness and Language

*First published Wed Jan 16, 2008*

The philosophical debate over innate ideas and their role in the acquisition of knowledge has a venerable history. It is thus surprising that very little attention was paid until early last century to the questions of how *linguistic* knowledge is acquired and what role, if any, innate ideas might play in that process.

To be sure, many theorists have recognized the crucial part played by language in our lives, and have speculated about the (syntactic and/or semantic) properties of language that enable it to play that role. However, few had much to say about the *properties of us* in virtue of which we can learn and use a natural language. To the extent that philosophers before the 20th century dealt with language acquisition at all, they tended to see it as a product of our general ability to reason — an ability that makes us special, and that sets us apart from other animals, but that is not tailored for language learning in particular.

In Part 5 of the *Discourse on the Method*, for instance, Descartes identifies the ability to use language as one of two features distinguishing people from "machines" or "beasts" and speculates that even the stupidest people can learn a language (when not even the smartest beast can do so) because human beings have a "rational soul" and beasts "have no intelligence at all." (Descartes 1984: 140-1.) Like other great philosopher-psychologists of the past, Descartes seems to have regarded our acquisition of concepts and knowledge ('ideas') as the main psychological mystery, taking language acquisition to be a relatively trivial matter in comparison; as he puts it, albeit ironically, "it patently requires very little reason to be able to speak." (1984: 140.)

All this changed in the early twentieth century, when linguists, psychologists, and philosophers began to look more closely at the phenomena of language learning and mastery. With advances in syntax and semantics came the realization that knowing a language was not merely a matter of associating words with concepts. It also crucially involves knowledge of *how to put words together,* for it's typically sentences that we use to express our thoughts, not words in isolation.

If that's the case, though, language mastery can be no simple matter. Modern linguistic theories have shown that human languages are vastly complex objects. The syntactic rules governing sentence formation and the semantic rules governing the assignment of meanings to sentences and phrases are immensely complicated, yet language users apparently apply them hundreds or thousands of times a day, quite effortlessly and unconsciously. But if knowing a language is a matter of knowing all these obscure rules, then acquiring a language emerges as the monumental task of *learning* them all. Thus arose the question that has driven much of modern linguistic theory: How could mere children learn the myriad intricate rules that govern linguistic expression and comprehension in their language — and learn them solely from exposure to the language spoken around them?

Clearly, there is something very special about the brains of human beings that enables them to master a natural language — a feat usually more or less completed by age 8 or so. §2.1 of this article introduces the idea, most closely associated with the work of the MIT linguist Noam Chomsky, that what is special about human brains is that they contain a specialized 'language organ,' an innate mental 'module' or 'faculty,' that is dedicated to the task of mastering a language.

On Chomsky's view, the language faculty contains innate knowledge of various linguistic rules, constraints and principles; this innate knowledge constitutes the 'initial state' of the language faculty. In interaction with one's experiences of language during childhood — that is, with one's exposure to what Chomsky calls the 'primary

linguistic data' or '*pld*' (see §2.1) — it gives rise to a new body of linguistic knowledge, namely, knowledge of a specific language (like Chinese or English). This 'attained' or 'final' state of the language faculty constitutes one's 'linguistic competence' and includes knowledge of the grammar of one's language. This knowledge, according to Chomsky, is essential to our ability to speak and understand a language (although, of course, it is not sufficient for this ability: much additional knowledge is brought to bear in 'linguistic performance,' that is, actual language use).[1]

§§2.2-2.5 discuss the main arguments used by Chomsky and others to support this 'nativist' view that what makes language acquisition possible is the fact that much of our linguistic knowledge is unlearned; it is innate or inborn, part of the initial state of the language faculty.[2] Section 3 presents a number of other avenues of research that have been argued to bear on the innateness of language, and shows how recent empirical research about language learning and the brain may challenge the nativist position. Because much of this material is very new, and because my conclusions (many of which are tentative) are highly controversial, more references to the empirical literature than are normal in an encyclopedia article are included. The reader is encouraged to follow up on the research cited and assess the plausibility of linguistic nativism for him or herself: whether language is innate or not is, after all, an empirical issue.

# 1. Chomsky's Case against Skinner

The behaviorist psychologist B.F. Skinner was the first theorist to propose a fully fledged theory of language acquisition in his book, *Verbal Behavior* (Skinner 1957). His theory of learning was closely related to his theory of linguistic behavior itself. He argued that human linguistic behavior (that is, our own utterances and our responses to the utterances of others) is determined by two factors: (i) the current features of the environment impinging on the speaker, and (ii) the speaker's history of reinforcement (i.e., the giving or withholding of rewards and/or punishments in response to previous linguistic behaviors). Eschewing talk of the mental as unscientific, Skinner argued that 'knowing' a language is really just a matter of having a certain set of behavioral

dispositions: dispositions to say (and do) appropriate things in response to the world and the utterances of others. Thus, knowing English is, in small part, a matter of being disposed to utter "Please close the door!" when one is cold as a result of a draught from an open door, and of being disposed (other things being equal) to utter "OK" and go shut a door in response to someone else's utterance of that formula.

Given his view that knowing a language is just a matter of having a certain set of behavioral dispositions, Skinner believed that learning a language just amounts to acquiring that set of dispositions. He argued that this occurs through a process that he called *operant conditioning*. ('Operants' are behaviors that have no discernible law-like relation to particular environmental conditions or 'eliciting stimuli.' They are to be contrasted with 'respondents,' which are reliable or reflex responses to particular stimuli. Thus, blinking when someone pokes at your eye is a respondent; episodes of infant babbling are operants.) Skinner held that most human verbal behaviors are operants: they start off unconnected with any particular stimuli. However, they can acquire connections to stimuli (or other behaviors) as a result of conditioning. In conditioning, the behavior in question is made more (or in some paradigms less) likely to occur in response to a given environmental cue by the imposition of an appropriate 'schedule of reinforcement': rewards or punishments are given or withheld as the subject's response to the cue varies over time.

According to Skinner, language is learned when children's verbal operants are brought under the 'control' of environmental conditions as a result of training by their caregivers. They are rewarded (by, e.g., parental approval) or punished (by, say, a failure of comprehension) for their various linguistic productions and as a result, their dispositions to verbal behavior gradually converge on those of the wider language community. Likewise, Skinner held, 'understanding' the utterances of others is a matter of being trained to perform appropriate behaviors in response to them: one understands 'Shut the door!' to the extent that one responds appropriately to that utterance.

In his famous review of Skinner's book, Chomsky (1959) effectively demolishes Skinner's theories of both language mastery and language learning. First, Chomsky argued, mastery of a language is not merely a matter of having one's verbal behaviors 'controlled' by various elements of the environment, including others' utterances. For language use is (i) stimulus independent and (ii) historically unbound. Language use is *stimulus independent*: virtually any words can be spoken in response to any environmental stimulus, depending on one's state of mind. Language use is also *historically unbound*: what we say is not determined by our history of reinforcement, as is clear from the fact that we can and do say things that we have not been trained to say.

The same points apply to comprehension. We can understand sentences we have never heard before, even when they are spoken in odd or unexpected situations. And how we react to the utterances of others is again dependent largely on our state of mind at the time, rather than any past history of training. There are linguistic conventions in abundance, to be sure, but as Chomsky rightly pointed out, human 'verbal behavior' is quite disanalogous to a pigeon's disk-pecking or a rat's maze-running.. Mastery of language is not a matter of having a bunch of mere behavioral dispositions. Instead, it involves a wealth of pragmatic, semantic and syntactic knowledge. What we say in a given circumstance, and how we respond to what others say, is the result of a complex interaction between our history, our beliefs about our current situation, our desires, *and our knowledge of how our language works*. Skinner's first big mistake, then, was in failing to recognize that language mastery involves knowledge (or, as Chomsky later called it 'cognizance') of linguistic rules and conventions.

His second big mistake was related to this one: he failed to recognize that acquiring mastery of a language is not a matter of being trained what to say. It's simply false, says Chomsky, that "a careful arrangement of contingencies of reinforcement by the verbal community is a necessary condition of language learning." (1959:39) First, children learning language do not appear to be being 'conditioned' at all! Explicit training (such as a dog receives when learning to bark on command) is simply not a feature of language acquisition. It's only comparatively rarely that parents correct (or explicitly reward) their children's linguistic sorties; children learn much of what they know about language from watching TV or passively listening to adults; immigrant children learn a second language to native speaker fluency in the school playground; and even very young children are capable of linguistic innovation, saying things undreamt of by their parents. As Chomsky concludes: "It is simply not true that children can learn language only through 'meticulous care' on the part of adults who shape their verbal repertoire through careful differential reinforcement." (1959:42)

Secondly, Chomsky argued — and here we see his first invocation of the famous 'poverty of the stimulus' argument, to be discussed in more detail in §2.2 below — it is unclear that conditioning could *even in principle* give rise to a set of dispositions rich enough to generate the full range of a person's linguistic behavior. In order, for example, to acquire the appropriate set of dispositions concerning the word *car*, one would have to be trained on vast numbers of sentences containing that word: one would have to hear *car* in object position and *car* in subject position; *car* modified by adjectives and *car* unmodified; *car* embedded in opaque contexts (e.g. in propositional attitude ascriptions) and *car* used transparently; and so on. But the 'primary linguistic data,' usually referred to as the '*pld*' and comprising the set of sentences to which a child is exposed during language learning (plus any analysis performed by the child on those sentences; see below), simply cannot be assumed to contain enough of these 'minimally differing sentences' to fully determine a person's dispositions with respect to that word. Instead, Chomsky argued, what determines one's dispositions to use *car* is one's knowledge of that word's syntactic and semantic properties (e.g., *car* is a noun referring to cars), together with one's knowledge of how elements with those properties function in the language as a whole. So even if language mastery were (in part) a matter of having dispositions concerning *car*, the mechanism of conditioning would be unable to give rise to them. The training set to which children have access is simply too limited: it doesn't contain enough of the right sorts of exemplars.

In sum: Skinner was mistaken on all counts. Language mastery is not merely a matter of having a set of bare behavioral dispositions. Instead, it involves intricate and detailed knowledge of the properties of one's language. And language learning is not a matter of being trained what to say. Instead, children learn language just from hearing it spoken around them, and they learn it effortlessly, rapidly, and without much in the way of overt instruction.

These insights were to drive linguistic theorizing for the next fifty years, and it's worth emphasizing just how radical and exciting they were at the time. First, the idea that explaining language use involves attributing *knowledge* to speakers flouted the prevailing behaviorist view that talking about mental states was unscientific because mental states are unobservable. It also raised several pressing empirical question that linguists are still debating. For example, what is the *content* of speakers' knowledge of language?[3] What sorts of facts about language are represented in speakers' heads? And how does this knowledge actually function in the psychological processes of language production and comprehension: what are the *mechanisms* of language use?

Secondly, the idea that children learn language essentially on their own was a radical challenge to the prevailing behaviorist idea that all learning involves reinforcement. In addition, it made clear our need for a more 'cognitive' or 'mentalistic' conception of how language learning occurs, and vividly raised the question — our focus in this article — of what might be the *preconditions* for that process. As we will see in the next section, Chomsky was ready with a theory addressing each of these points.

# 2. Arguments for the Innateness of Language

## 2.1 What do Children Learn when they Learn Language?

At the same time as the behaviorist program in psychology was waning under pressure from Chomsky and others, linguists were abandoning what is known as 'American Structuralism' in the theory of syntax. Like the behaviorists, the structuralists (e.g., Harris, 1951) refused to postulate irreducibly theoretical entities; they insisted that syntactic categories (such as 'noun phrase' ('NP') or 'verb phrase' ('VP'), etc.) be reducible to properties of actual utterances (collected in 'corpora' — lists of things people have said). In his landmark book, *Syntactic Structures* (1957), however, Chomsky argued that because corpora can contain only finitely many sentences, no attempt at reduction can succeed. Linguists need theoretical constructs that capture regularities going beyond the set of actual utterances, and that allow them to predict the properties of novel utterances. But if the category NP, for instance, is to include noun phrases that haven't been uttered yet, the meaning of *noun phrase* can't be exhausted by what's in the corpus: the structuralists' positivistic strictures on theoretical kinds are misguided.

In addition, the structuralists had attempted to capture the syntactic properties of languages in terms of simple rewrite rules known as 'phrase structure rules.' Phrase structure rules describe the internal syntactic structures of sentence types; interpreted as rewrite rules, they can be used to generate or construct sentences. Thus, the rule S → NP VP, for instance, says that a sentence symbol S can be rewritten as the symbol NP followed by the symbol VP, and tells you that a sentence consists of a noun phrase followed by a verb phrase. (This information can be represented via a tree-diagram, as in Fig. 1a, or by a phrasemarker (or labeled bracketing), as in Fig. 1b.)



$$[[\underline{\quad}]_{NP}[\underline{\quad}]_{VP}]_S$$

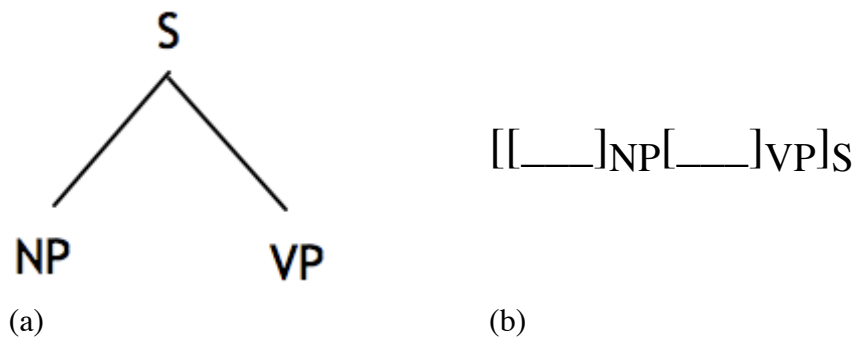(a)                                                        (b)

Figure 1. Phrasemarkers representing a sentence as consisting of a noun phrase and a verb phrase via (a) a tree diagram or (b) a labeled bracketing.

Other rules, (such as NP → Det N, VP → V NP, Det → *a, the, …*, etc., V→ *hit, kiss…*, etc.; N → *boy, girl,…*, etc.) are subsequently applied, and (with still further rules not discussed here) allow for the generation of sentences such as *The boy kissed the girl, The girl hits the boy*, and so on.

Chomsky argued (on technical grounds; see Chomsky 1957, ch.1) that grammars must be enriched with a second type of rule, known as 'transformations.' Unlike phrase structure rules, transformations operate on whole sentences (or more strictly, their phrasemarkers); they allow for the generation of new sentences (/phrasemarkers) out of old ones. The *Passive* transformation described in Chomsky 1957:112, for instance, specifies how to turn an active sentence (/phrasemarker) into a passive one. Simplifying somewhat, you take an active phrasemarker of the form $NP - Aux - V - NP$, like *Kate is biting Mark,* and rearrange its elements $x_1 - x_2 - x_3 - x_4$ as follows: $x_4 - x_2 + be + en - x_3 + by - x_1$ to get *Mark bite (+ is + en) by Kate*. The parenthetical + *en* and + *is* invoke further operations on the verb *bite* that transform it into *is being bitten,* and ultimately *Kate is biting Mark* is 'transformed' into *Mark is being bitten by Kate*.

Only a grammar containing both phrase structure and transformation rules, Chomsky argued, could generate a natural language — 'generate' in the sense that by stepwise application of the rules, one could in principle build up from scratch all and only the sentences that the language contains. Hence, Chomsky urged the development of generative grammars of this type.

Syntactic theory has now gone well beyond this early vision — both phrase structure and transformation rules were abandoned in successive linguistic revolutions wrought by Chomsky and his students and colleagues (see Newmeyer 1986, 1997 for a history of generative linguistics).

But what has not changed — and what is important for our purposes — is that in every version of the grammar of (say) English, the rules governing the syntactic structure of sentences and phrases are stated in terms of syntactic categories that are highly abstracted from the properties of utterances that are accessible to experience. As an example of this, consider the notion of a *trace*. Traces are symbols that appear in phrasemarkers and mark the path of an element as it is moved from one position to another at various stages of a sentence's derivation, as in (1), where $t_i$ markes the NP *Jacob*'s position at an earlier stage in the derivation.

    1. Jacob$_i$ seems [$t_i$ to have vanished]

But while traces are vital to the statement of many syntactic rules and regularities, they are 'empty categories' — they are not audible in the sentence as spoken. (See Chomsky 1981 and Lasnik and Uriagereka 1986 for more on

traces and other empty categories.) Traces (and other similarly abstract properties of languages) thus raise a question for the theory of language acquisition. For if, as Chomsky maintains, mastery of language involves knowledge of rules stated in terms of sentences' syntactic properties, and if those properties are not so to speak 'present' in the data, but are rather highly abstract and 'unobservable,' then it becomes hard to see how children could possibly acquire knowledge of the rules concerning them. As a consequence, children's feat in learning a language appears miraculous: how could a child learn the myriad rules governing linguistic expression given only her exposure to the sentences spoken around her?[4]

In response to this question, most 20th century theorists followed Chomsky in holding that language acquisition could not occur unless much of the knowledge eventually attained were innate or inborn. The gap between what speaker-hearers know about language (its grammar, among other things) and the data they have access to during learning (the *pld*) is just too broad to be bridged by any process of learning alone. It follows that since children patently do learn language, they are not linguistic 'blank slates.' Instead, Chomsky and his followers maintained, human children are born knowing the 'Universal Grammar' or 'UG,' a theory describing the most fundamental properties of all natural languages (e.g., the facts that elements leave traces behind when they move, and that their movements are constrained in various ways). Learning a particular language thus becomes the comparatively simple matter of elaborating upon this antecedently possessed knowledge, and hence appears a much more tractable task for young children to attempt.

Over the years, two conceptions of the innate contribution to language learning and its elaboration during the learning process have been proposed. In earlier writings (e.g., Chomsky 1965), Chomsky saw learning a language as basically a matter of formulating and testing hypotheses about its grammar — unconsciously, of course. He argued that in order to acquire the correct grammar, the child must innately know a "a linguistic theory that specifies the form of the grammar of a possible human language" (1965:25) — she must know UG in other words. He saw this knowledge as being embodied in a suite of innate linguistic abilities, concepts, and constraints on the kinds of grammatical rules learners can propose for testing. On this view (1965:30-31), the inborn UG includes (i) a way of analyzing and representing the incoming linguistic data; (ii) a set of linguistic concepts with which to state grammatical hypotheses; (iii) a way of telling how the data bear on those hypotheses (an 'evaluation metric'); and (iv) a very restrictive set of constraints on the hypotheses that are available for consideration. (i) through (iv) constitute the 'initial state' of the language faculty, and the child arrives at the final state (knowledge of her language) by performing what is basically a kind of scientific inquiry into its nature.

By the 1980's, a less intellectualized conception of how language is acquired began to supplant the hypothesis-testing model. Whereas the early model saw the child as a 'little scientist,' actively (if unconsciously) figuring out the rules of grammar, the new 'parameter-setting' model conceived language acquisition as a kind of growth or maturation; language acquisition is something that happens to you, not something you do. The innate UG was no longer viewed as a set of tools for inference; rather, it was conceived as a highly articulated set of representations of actual grammatical principles. Of course, since not everyone ends up speaking the same language, these innate representations must allow for some variation. This is achieved in this model *via* the notion of a 'parameter': some of the innately represented grammatical principles contain variables that may take one of a certain (highly restricted) range of values. These different 'parameter settings' are determined by the child's linguistic experience, and result in the acquisition of different languages. Thus, Chomsky (1988:61-62) compared the learner to a switchbox: just as a switchbox's circuitry is all in place but for some switches that need to be flicked to one position or another, the learner's knowledge of language is basically all in place, but for some linguistic 'switches' that are set by linguistic experience.

To illustrate how parameter setting works, consider a simplified example (discussed in more detail in Chomsky 1990:644-45). All languages require that sentences have subjects, but whereas some languages (like English) require that the subject be overt in the utterance, other languages (like Spanish) allow you to leave the subject out of the sentence when it is written or spoken. Thus, a Spanish speaker who wanted to say that he speaks Spanish could say *Hablo español* (leaving out the first personal pronoun *yo*) without violating the rules of Spanish, whereas an English speaker wanting to express that thought could not say *Speak Spanish* without violating the rules of English: to speak grammatically, he must say *I speak Spanish*. The parameter-setting model

accommodates this sort of difference by proposing that there is a 'Null Subject Parameter,' which is set differently in English and Spanish speakers: Spanish speakers set it to 'Subject Optional,' whereas in English speakers, it is set to 'Subject Obligatory.' How? One proposal is that the parameter is set by default to 'Subject Obligatory' and that hearing a subjectless sentence causes it to be set to 'Subject Optional.' Since children learning Spanish frequently hear subjectless sentences, whereas those learning English do not, the parameter setting is switched in the Spanish learner, but remains set at the default for the English learner. (Roeper and Williams 1987 is the *locus classicus* for parameter-setting models; Ayoun 2003 is more up-to-date; Pinker, 1997: ch.3 provides a helpful, non-technical overview.)

These two approaches to language acquisition clearly differ significantly in their conception of the nature of the learning process and the learner's role in it, but we are not concerned to evaluate their respective merits here. Rather, the important point for our purposes is that they both attribute substantial amounts of innate information *about language* to the language learner. In what follows, we will look in more detail at the various arguments that have been used to support this 'nativist' theory of language acquisition. We will focus on the following question:

> What evidence is there that children come to the language learning task equipped with a specialized store of inborn linguistic information, such as that specified in the linguist's theory of Universal Grammar?

Terminological Note: As Chomsky acknowledges (e.g., 1986:28-29), 'Universal Grammar' is used with a systematic ambiguity in his writings. Sometimes, the term refers to the inborn knowledge of language that learners are hypothesized to possess — the content of the 'initial state' of the language faculty — *whatever that knowledge (/content) turns out to be*. Other times, 'Universal Grammar' is used to refer to certain *specific proposals* as to the content of our innate linguistic knowledge, such as the Government-Binding theorist's claim that we have inborn knowledge of such things as the Principle of Structure Dependence, Binding theory, Theta theory, the Empty Category Principle, etc.

This ambiguity is important when one is evaluating Chomskyan claims that we have innate knowledge of UG. For on the first reading of 'Universal Grammar' distinguished above, that claim will be true so long as *any* form of nativism turns out to be true of language learners (i.e., so long as they possess *any* inborn knowledge about language). On the second reading, however, it is possible that learners have innate knowledge of language *without* that knowledge's being knowledge of UG (as currently described by linguists): learners might know things about language, yet *not* know Binding Theory, or the Principle of Structure Dependence, etc.

In this entry, 'Universal Grammar' will always be used in the second of these senses, to refer to a specific theory as to the content of learners' innate knowledge of language. Where the issue concerns merely their having *some or other* innate knowledge about language (and is neutral on the question of whether any particular theory about that knowledge is true), I will talk of 'innate linguistic information.' Clearly, an argument to the effect that speakers have inborn knowledge of UG entails the claim that they have innate linguistic information at their disposal. The reverse, however, is not the case: there might be reason to think that a speaker knows something about language innately, without its constituting reason to think that what they know is Universal Grammar as described by Chomksyan linguists; Chomksy might be right that we have innate knowledge about language, but wrong about what the *content* of that knowledge is. These issues will be clarified, as necessary, below.

## 2.2 Chomsky's 'Poverty of the Stimulus' Argument for the Innateness of Language

As we saw in §1.1, one of the conclusions Chomsky drew from his (1959) critique of the Skinnerian program was that language cannot be learned by mere association of ideas (such as occurs in conditioning). Since language mastery involves knowledge of grammar, and since grammatical rules are defined over properties of utterances that are not accessible to experience, language learning must be more like theory-building in science. Children appear to be 'little linguists,' making highly theoretical hypotheses about the grammar of their language and testing them against the data provided by what others say (and do):

> It seems plain that language acquisition is based on the child's discovery of what from a formal point of view is a deep and abstract theory — a generative grammar of his language — many of the concepts and principles of that are only remotely related to experience by long and intricate chains of quasi-inferential steps. (Chomsky 1965:58)

However, argued Chomsky, just as conditioning was too weak a learning strategy to account for children's ability to acquire language, so too is the kind of inductive inference or hypothesis-testing that goes on in science. Successful scientific theory-building requires huge amounts of *data,* both to suggest plausible-seeming hypotheses and to weed out any false ones. But the data children have access to during their years of language learning (the 'primary linguistic data' or '*pld'*) are highly impoverished, in two important ways:

  i. they constitute a small finite sample of the infinitely many sentences natural languages contain
  ii. they do not reliably contain the kinds of sentences that learners need to falsify incorrect hypotheses

The first type of inadequacy is, of course, endemic to any kind of empirical inquiry: it is simply the problem of the underdetermination of theories by their evidence. Cowie has argued elsewhere that underdetermination per se cannot be taken to be evidence for nativism: if it were, we would have to be nativists about *everything* that people learn (Cowie 1994; 1999). What of the second kind of impoverishment? If the evidence about language available to children does not enable them to reject false hypotheses, and if they nonetheless hit on the correct grammar, then language learning could not be a kind of scientific inquiry, which depends in part on being able to find evidence to weed out incorrect theories. And indeed, this is what Chomsky argues: since the *pld* are not sufficiently rich or varied to enable a learner to arrive at the correct hypothesis about the grammar of the language she is learning, language could not be learned from the *pld*.

For consider: The fact (i) that the *pld* are finite whereas natural languages are infinite shows that children must be generalizing beyond the data when they are learning their language's grammar: they must be proposing rules that cover as-yet unheard utterances. This, however, opens up room for error. In order to recover from particular sorts of error, children would need access to particular kinds of data. If those data don't exist, as (ii) asserts, then children would not be able to correct their mistakes. Thus, since children do eventually converge on the correct grammar for their language, they mustn't be making those sorts of errors in the first place: something must be stopping them from making generalizations that they cannot correct on the basis of the *pld*.

Chomsky (e.g., 1965: 30-31) expresses this last point in terms of the need for *constraints* — on grammatical concepts, on the hypothesis space, on the interpretation of data — and proposes that it is innate knowledge of UG that supplies the needed limitations. On this view, children learning language are not open-minded or naïve theory generators — they are not 'little scientists.' Instead, the human language-learning mechanism (the 'language acquisition device' or 'LAD') embodies built-in knowledge about human languages, knowledge that prevents learners from entertaining most possible grammatical theories. As Chomsky puts it:

> A consideration of…the degenerate quality and narrowly limited extent of the available data … leave[s] little hope that much of the structure of the language can be learned by an organism initially uninformed as to its general character. (1965:58)

Chomsky rarely states the argument from the poverty of the stimulus in its general form, as Cowie has done here. Instead, he typically presents it *via* an example. One of these concerns learning how to form 'polar interrogatives,' i.e., questions demanding *yes* or *no* by way of answer, *via* a mechanism known as 'auxiliary fronting.'[5] Suppose that a child heard pairs of sentences like the following:

  1a. Jacob is happy today
  1b. Is Jacob happy today?
  2a. The girls are dancing
  2b. Are the girls dancing?

She wants to figure out the rule you use to turn declaratives like (1a) and (2a) into interrogatives like (1b) and (2b). Here are two possibilities:

$H_1$. Find the first occurrence of *is* in the sentence and move it to the front.

$H_2$. Find the first occurrence of *is* following the subject nounphrase ('NP') of the sentence, and move it to the front.

Both hypotheses are adequate to account for the data the learner has so far encountered. To any unbiased scientist, though, $H_1$ would surely appear preferable to $H_2$, for it is *simpler* — it is shorter, for one thing, and does not refer to theoretical properties, like being a NP, being instead formulated in terms of 'observable' properties like word order. Nonetheless, $H_1$ is false, as is evident when you look at examples like (3):

3a. [The girl who is in the jumping castle]$_{NP}$ is Kayley's daughter

3b. *Is [the girl who in the jumping castle]$_{NP}$ is Kayley's daughter?

3c. Is [the girl who is in the jumping castle]$_{NP}$ Kayley's daughter?

$H_1$ generates the ungrammatical question (3b), whereas $H_2$ generates the correct version, (3c).[6] Now, you and I and every other English speaker know (in some sense — see §3.2.1a) that $H_1$ is false and $H_2$ is correct. *That* we know this is evident, Chomsky argues, from the fact that we all know that (3b) is not the right way to say (3c). The question is how we could have learnt this.

Suppose, for example, that based on her experience of (1) and (2), a child were to adopt $H_1$. How would she discover her error? There would seem to be two ways to do this. First, she could use $H_1$ in her own speech, utter a sentence like (3b), and be corrected by her parents or caregivers; second, she could hear a sentence like (3c) uttered by a competent speaker, and realize that that sentence is not generated by her hypothesis, $H_1$. But typically parents don't correct their children's ill formed utterances (see §2.2.1(c) for more on this), and worse, according to Chomsky, sentences like (3c) — sentences that are not generated by the incorrect rule $H_1$ and hence would falsify it— do not occur often enough in the *pld* to guarantee that every native English speaker will be able to get it right.

So in answer to the question: how do we learn that $H_2$ is better than $H_1$, Chomsky argued that we don't *learn* this at all! A better explanation of how we all know that $H_2$ is right and $H_1$ is wrong is that we were *born knowing* this fact. Or, more accurately, we were born knowing a certain principle of UG (the 'Principle of Structure Dependence'), which tells us that rules like $H_1$ are not worth pursuing, their ostensible 'simplicity' notwithstanding, and that we should always prefer rules, like $H_2$, which are stated in terms of sentences' *structural* properties. In sum, we know that $H_2$ is a better rule than $H_1$, but we didn't learn this from our experience of the language. Rather, this fact is a consequence of our inborn knowledge of UG.

Chomskyans contest that there are many other cases in which speaker-hearers know grammatical rules, the critical evidence in favor of which is missing from the *pld*. Kimball 1973:73-5, for instance, argues that complex auxiliary sequences like *might have been* are "vanishingly rare" in the *pld*, hence that children acquire competence with these constructions (in the sense of knowing the order in which to put the modal, perfect and progressive elements) without relevant experience. (Pullum and Scholz 2002, discuss two other well known examples.) Nativists thus conclude that numerous other principles of UG are innately known as well. Together, these UG principles place strong constraints on learners' linguistic theorizing, preventing them from making errors for which there are no falsifying data.

So endemic is the impoverishment of the *pld*, according to Chomskyans, that it began to seem as if the entire learning paradigm were inapplicable to language. As more and more and stricter and stricter innate constraints needed to be imposed on the learner's hypothesis space to account for their learning rules in the absence of relevant data, notions like hypothesis generation and testing seemed to have less and less purchase. This situation

fuelled the recent shift away from hypothesis testing models of language acquisition and towards parameter setting models discussed in §2.1 above.

## 2.2.1 Criticisms of the Poverty of the Stimulus Argument

Many, probably most theorists in modern linguistics and cognitive science have accepted Chomsky's poverty of the stimulus argument for the innateness of UG. As a result, a commitment to linguistic nativism has underpinned most research into language acquisition over the last 40-odd years. Nonetheless, it is important to understand what criticisms have been leveled against the argument, which I schematize as follows for convenience:

> The General Form of the Argument from the Poverty of the Stimulus
>
> 1. Mastery of a language consists (in part) of knowing its grammar.
> 2. In order to learn a certain rule of grammar, $G$ children would have to have access to certain sorts of data, $D$, which falsify competing hypotheses.
> 3. The primary linguistic data *(pld)* do not contain $D$.
>
> So
>
> 4. $G$ could not be learned.
> 5. This situation is quite general: many rules of grammar are unlearnable from the *pld*.
>
> So
>
> 6. UG is innately known.

**2.2.1(a) Premiss 1: Knowledge of grammar**

In the 1970's, philosophers contested Chomsky's use of the word 'know' to describe speakers' relations to grammar, arguing that unlike standard cases of propositional knowledge, most speakers are utterly unaware of grammatical rules (e.g., "Anaphors are bound, and pronominals and R-expressions are free in their binding domains") and many probably wouldn't understand them even if told what they are (Stich 1971). In response, Chomsky (e.g., 1980:92) began to use a technical term, 'cognize,' to describe the speaker-grammar relation, avoiding the philosophically loaded term, 'knowledge.'

However, while it is certainly legitimate to propose a special relationship between speakers and grammars, unanswered questions remain about the precise nature of cognizance. Is it a representational relation, like belief? If not, what does 'learning a grammar' amount to? If so, are speakers' representations of grammar 'explicit' or 'implicit' or 'tacit' — and what, exactly, do any of these terms mean? (See the papers collected in MacDonald 1995, for discussion of this last issue; see Devitt 2006 for arguments that there is no good reason to suppose that speakers use any representations of grammatical rules in their production and comprehension of language.) Relatedly, how does a speaker's cognizance of grammar (her 'competence,' in Chomskyan parlance) function in her linguistic 'performance' — i.e., in the actual production or comprehension of an utterance?

These issues bear on the argument from the poverty of the stimulus because that argument may appear more or less impressive depending on the answers one gives to them. If, for instance, one held that grammars are belief-like entities, explicitly represented in our heads in some internal code (cf. Stich 1978), then the question of how those beliefs are acquired and justified is indeed a pressing one — as, for different reasons, is the question of how they function in performance (see Harman 1967, 1969). However, if one were to deny that grammar is represented at all in the heads of speakers, like Devitt 2006 and Soames 1984, then the issue of how language is learned and what role 'evidence' etc. might play in that process takes on a very different cast. Or if, to take a third possibility, one were to reject generative syntax altogether and adopt a different conception of what the content of speakers' grammatical knowledge is — along the lines of Tomasello (2003), say — then that again

affects how one views the learning process. In other words, one's ideas about *what is learned* affect one's conception of *what is needed to learn it*. Less 'demanding' conceptions of the outputs of language acquisition require less demanding conceptions of its input (whether experiential or inborn); this last approach to the problem of language learning is discussed further in §2.2.1 below.

**2.2.1(b) Premiss 2: The learning algorithm**

In the example of polar interrogatives, discussed above, we saw how children apparently require explicit falsifying evidence in order to rule out the plausible-seeming but false hypothesis, $H_1$. Premiss 2 of the argument generalizes this claim: there are many instances in which learners need specific kinds of falsifying data to correct their mistakes (data that the argument goes on to assert are unavailable). These claims about the data learners would need in order to learn grammar are underpinned by certain assumptions about the learning algorithm they employ. For example, the idea that false hypotheses are rejected only when they are explicitly falsified in the data suggests that learners are incapable of taking any kind of probabilistic or holistic approach to confirmation and disconfirmation. Likewise, the idea that learners unequipped with inborn knowledge of UG are very likely indeed to entertain false hypotheses suggests that their method of generating hypotheses is insensitive to background information or past experience. (e.g., information about what sorts of generalizations have worked in other contexts

The non-nativist language learner as envisaged by Chomsky in the original version of the poverty of the stimulus argument, in other words, is limited to a kind of Popperian methodology — one that involves the enumeration of all possible grammatical hypotheses, each of which is tested against the data, and each of is rejected just in case it is explicitly falsified. As much work in philosophy of science over the last half century has indicated, though, nothing much of anything can be learned by this method: the world quite generally fails to supply falsifying evidence. Instead, hypothesis generation must be inductively based, and (dis)confirmation is a holistic matter.

Thus arise two problems for the Chomskyan argument. First, it is not all that surprising to discover that if language learners employed a method of conjecture and refutation, then language could not be learned from the data. In other words, the poverty of the stimulus argument doesn't tell us much we didn't know already. Secondly, and as a result, the argument is quite weak: it makes the negative point that language acquisition does not occur via a Popperian learning strategy, but it favors no specific alternative to this acquisition theory. In particular, the argument gives no more support to a nativist (UG-based) theory than to one that proposed (say) that learners formulate grammatical hypotheses based on their extraction of statistical information about the *pld* and that they may reject them for reasons other than outright falsification — because they lack explicit confirmation, or because they do not cohere with other parts of the grammar, for instance.

In reply, some Chomskyans (e.g., Matthews 2001) challenge non-nativists to produce these alternative theories and submit them to empirical test. It's pointless, they claim, for nativists to try to argue against theories that are mere gleams in the empiricist's eye, particularly when Chomsky's approach has been so fruitful and thus may be supported by a powerful inference to the best explanation. Others have argued explicitly against particular non-nativist theories — Marcus 1998, 2001, for instance, discusses the shortcomings of connectionist accounts of language acquisition.

A recent book by Michael Tomasello (Tomasello 2003) addresses the nativist's demand for an alternative theory directly. Tomasello argues that language learners acquire knowledge of syntax by using inductive, analogical and statistical learning methods, and by examining a broader range of data for the purposes of confirmation and disconfirmation. He argues that children formulate abstract syntactic generalizations rather late in the learning process (around the age of 4 or 5) and that their earliest utterances are governed by much less general rules of thumb, or 'constructions.' More abstract constructions, framed in increasingly adult-like and 'syntactic' terms, are progressively formulated through the application of pattern-recognition skills ('analogy') and a kind of statistical analysis of both incoming data and previously acquired constructions, which Tomasello calls 'functional distributional analysis.'[7]

Tomasello's theory differs from a Chomskyan approach in three important respects. First, and taking up a point mentioned in the previous section, it employs a different conception of linguistic competence, the end state of the learning process. Rather than thinking of competent speakers as representing the rules of grammar in the maximally abstract, simple and elegant format devised by generative linguists, Tomasello conceives of them as employing rules at a variety of different levels of abstraction, and, importantly, as employing rules that are not formulated in purely syntactic terms. He adopts a different type of grammar, called 'cognitive-functional grammar' or 'usage-based grammar,' in which rules are stated partly in terms of syntactic categories, but also in semantic terms, that is, in terms of their patterns of use and communicative function. A second respect in which Tomasello's approach differs from that of most theorists in the Chomskyan tradition, is in employing a much richer conception of the 'primary linguistic data,' or *pld*. For generative linguists, the *pld* comprises a set of sentences, perhaps subject to some preliminary syntactic analysis, and the child learning grammar is thought of as embodying a function which maps that set of sentences onto the generative grammar for her language. On Tomasello's conception, the *pld* includes not just a set of sentences, but also facts about how sentences are used by speakers to fulfill their communicative intentions. On his view, semantic and contextual information is also used by children for the purposes of acquiring grammatical knowledge.

Tomasello argues that by adopting a more 'user-friendly' conception of natural language grammars and by radically expanding one's conception of the language-relevant information available to children learning language, the 'gap' exploited by the argument from the poverty of the stimulus — that is, the gap between what we know about language and the data we learn it from — in large part disappears. This gives rise to a third important respect in which Tomasello's theory differs from that of the linguistic nativist. On his view, children learn language without the aid of any inborn linguistic information: what children bring to the language learning task — their innate endowment — is not language-specific. Instead, it consists of 'mind reading,' together with perceptual and cognitive skills that are employed in other domains as well as language learning. These skills include: (i) the ability to share attention with others; (ii) the ability to discern others' intentions (including their communicative intentions); (iii) the perceptual ability to segment the speech stream into identifiable units at different levels of abstraction; and (iv) general reasoning skills, such as the ability to recognize patterns of various sorts in the world, the ability to make analogies between patterns that are similar in certain respects, and the ability to perform certain sorts of statistical analysis of these patterns. Thus, Tomasello's theory contrasts strongly with the nativist approach.

Although assessing Tomasello's theory of language acquisition is beyond the scope of this entry, this much can be said: the oft-repeated charge that empiricists have failed to provide comprehensive, testable alternatives to Chomskyanism is no longer sustainable, and if the what and how of language acquisition are along the lines that Tomasello describes, then the motivation for linguistic nativism largely disappears.

**2.2.1(c) Premiss 3: What do the *pld* contain?**

A third problem with the poverty of the stimulus argument is that there has been little systematic attempt to provide empirical evidence supporting its assertions about what the *pld* contain. This is an old complaint (cf. Sampson 1989) which has recently been renewed with some vigor by Pullum and Scholz 2002, Scholz and Pullum 2002, and Sampson 2002. Pullum and Scholz provide evidence that, contrary to what Chomsky asserts in his discussion of polar interrogatives, children can expect to encounter plenty of data that would alert them to the falsity of $H_1$. Sampson 2002, mines the 'British National Corpus/demographic,' a 100 million word corpus of everyday British speech (available online at http://info.ox.ac.uk/bnc/), for evidence that contrary to Kimball's contention that complex auxiliaries are 'vanishingly rare,' they in fact occur quite frequently (somewhere from once every 10,000 words to once every 70,000 words, or once every couple of days to once a week).

Chomskyans respond in two main ways to findings like this. First, they argue, it is not enough to show that *some* children can be expected to hear sentences like *Is the girl in the jumping castle Kayley's daughter? All* children learn the correct rule, so the claim must be that *all* children are guaranteed to hear sentences of this form — and this claim is still implausible, data like those just discussed notwithstanding.[8] In order to take this question further, it would be necessary to determine when in fact children master the relevant structures, and vanishingly little work has been done on this topic. Sampson 2002:82ff. found no well-formed auxiliary fronted questions

(like *Is the girl who is in the jumping castle Kayley's daughter?*) in his sample of the British National Corpus. He notes that in addition to supporting Chomsky's claims about the poverty of the *pld*, such data simultaneously problematize his claims about children's knowledge of the auxiliary-fronting rule itself. Sampson found that speakers invariably made errors when apparently attempting to produce complex auxiliary-fronted questions, and often emended their utterance to a tag form instead (*e.g., The girl who's in the jumping castle is Kayley's daughter, isn't she?*). Hespeculates that the construction is not idiomatic even in adult language, and that speakers learn to form and decode such questions much later in life, after encountering them in written English. If that were the case, then the lack of complex auxiliary fronted questions in the *pld* would be both unsurprising and unproblematic: young children don't hear the sentences, but nor do they learn the rule. To my knowledge, children's competence with the auxiliary fronting rule has not been addressed empirically.[9]

Secondly, Chomskyans may produce other versions of the poverty of the stimulus argument. For instance, Crain 1991 constructs a poverty of the stimulus argument concerning children's acquisition of knowledge of certain constraints on movement. However, while Crain's argument carefully documents children's conformity to the relevant grammatical rules, its nativist conclusion still relies on unsubstantiated intuitions as to the non-occurrence of relevant forms or evidence in the *pld*. It is thus inconclusive. (Cf. Crain 1991; Crain's experiments and their implications are discussed in Cowie 1999 ; Cf. also Crain and Pietrowski 2001, 2002).

**2.2.1(d) The validity of the argument**

The argument from (1), (2), and (3) to (4) appears valid. However, as is implicit in my discussion of premiss (2), an equivocation between different senses of 'learning' threatens. What (1)-(3) show, if true, is that grammar G can't be learned from the *pld* by a learner *using a 'Popperian' learning strategy,* that is, a strategy of 'bold conjecture' and refutation. What (4) concludes, however, is that G is unlearnable, period, from the *pld* — a move that several authors authors, particularly connectionists, have objected to. (See especially Elman *et al*. 1996 and Elman 1998 for criticisms of Chomskyan nativism along these lines; see Marcus 1998 and 2001 for responses.)

Chomskyans typically take this point, conceding that the argument from the poverty of the stimulus is not apodeictic. Nonetheless, they claim, it's a very good argument, and the burden of proof belongs with their critics. After all, nativists have shown the falsity of the only non-nativist acquisition theories that are well-enough worked out to be empirically testable, namely, Skinnerian behaviorism and Popperian conjecture and refutation. In addition, they have proposed an alternative theory, Chomskyan nativism, which is more than adequate to account for the phenomena. In empirical science, this is all that they can reasonably be required to do. The fact that there might be other *possible* acquisition algorithms which *might* account for children's ability to learn language is neither here nor there; nativists are not required to argue against mere possibilities.

In response, some non-nativists have argued that UG-based theories are not in fact good theories of language acquisition. Tomasello (2003: 182ff.), for instance, identifies two major areas of difficulty for UG-based theories, such as the principles-and-parameters approach. First, there is the 'linking' problem, deriving from the fact of linguistic diversity: almost no UG-based accounts explain how children link the highly abstract categories of UG to their instantiations in the particular language they happen to be learning.[10] His example is the category 'Head,' In order to set the 'Head parameter,' a child needs to be able to identify which words in the stream of noise she is hearing are in fact clausal heads. But heads "do not come with identifying tags on them in particular languages; they share no perceptual features in common across languages, and so their means of identification cannot be specified in [UG]" (Tomasello 2003:183). Second, there is the problem of developmental change, also emphasized by Sokolov and Snow, 1991. It is difficult to see how UG-based approaches can account for the fact that children's linguistic performance seems to emerge piecemeal over time, rather than emerging in adult-like form all at once, as the parameter-setting model suggests it should.[11] In response, generativists have appealed to such notions as 'maturational factors' or 'performance factors.' But, Tomasello argues, such measures are ad hoc in the absence of a detailed specification of what these maturational or performance factors are, and how they give rise to children's actual performance.

At the very least, such objections serve to equalize the burden of proof: non-nativists certainly have work to do, but so too do nativists. Merely positing an innate UG and a 'triggering' mechanism by which it 'grows' into full-fledged language is insufficient. Nativists need to show how their theory can account for the known course of language acquisition. Merely pointing out that there is a *possibility* that such theories are true, and that they would, if true, explain how language learning occurs in the face of an allegedly impoverished stimulus, is only part of the job.

### 2.2.1(e) Premiss 5: How general is the poverty of the stimulus?

Because they are defending the view that all of UG is inborn, Chomskyans must be credited with holding that the primary data are impoverished quite generally. That is, if the innateness of UG *tout court* is to be supported by poverty of the stimulus considerations, the idea must be that the cases that nativists discuss in detail (polar interrogatives, complex auxiliaries, etc.) are but the tip of the unlearnable iceberg. Nativists quite reasonably do not attempt to defend this claim by endless enumeration of cases. Rather, they turn to another kind of argument to support the 'global impoverishment' position. This argument is sometimes called the 'Logical Problem of Language Acquisition'; here, we will call it 'The Unlearning Problem.' It will be discussed in section 3.

### 2.2.1(f) The validity of the argument (II): What is inborn?

Suppose that the primary linguistic data were impoverished in all the ways that nativists claim and suppose, too, that children know a bunch of things for which there is no evidence available — suppose, as Hornstein and Lightfoot (1981:9) put it, that "[p]eople attain knowledge of the structure of their language for which *no* evidence is available in the data to which they are exposed as children." What follows from this is that there must be constraints on the learning mechanism: children do not enumerate all possible grammatical hypotheses and test them against the data. Some possible hypotheses must be ruled out *a priori*. But, critics allege, what does *not* follow from this is *any particular view about the nature of the requisite constraints*. (Cowie 1999: ch.8.) A fortiori, what does not follow from this is the view that Universal Grammar (construed as a theory about the structural properties common to all natural languages, per Terminological Note 2 above) is inborn.

For all the poverty of the stimulus argument shows, the constraints in question might indeed be language-specific and innate, but with contents quite different from those proposed in current theories of UG. Or, the constraints might be innate, but not language-specific. For instance, as Tomasello 2003 argues, children's early linguistic theorizing appears to be constrained by their inborn abilities to share attention with others and to discern others' communicative intentions. On his view, a child's early linguistic hypotheses are based on the assumption that the person talking to him is attempting to convey information about the thing(s) that they are both currently attending to. (Another example of an innate but non-language specific constraint on language learning derives from the structure of the mammalian auditory system; 'categorical perception,' and is relation to the acquisition of phonological knowledge is discussed below, §3.3.4.). Another alternative is that the constraints might be learned, that is, derived from past experiences. An example again comes from Tomasello (2003). He argues that *entrenchment*, or the frequency with which a linguistic element has been used with a certain communicative function, is an important constraint on the development of children's later syntactic knowledge. For instance, it has been shown experimentally that the more often a child hears an element used for a particular communicative purpose, the less likely she is to extend that element to new contexts. (See Tomasello 2003:179).

In short, there are many ways to constrain learners' hypotheses about how their language works. Since the poverty of the stimulus argument merely indicates the need for constraints, it does not speak to the question of what sorts of constraints those might be.

In response to this kind of point, Chomskyans point out that the innateness of UG is an empirical hypothesis supported by a perfectly respectable inference to the best explanation. Of course there is a *logical* space between the conclusion that something constrains the acquisition mechanism and the Chomskyan view that these constraints are inborn representations of Binding Theory, Theta theory, the ECP, the principle of Greed or Shortest Path and so on. But the mere fact that the argument from the poverty of the stimulus doesn't *prove* that UG is innately known is hardly reason to complain. This is science, after all, and demonstrative proofs are

neither possible nor required. What the argument from the poverty of the stimulus provides is good reason to think that there are strong constraints on the learning mechanism. UG is at hand to supply a theory of those constraints. Moreover, that theory has been highly productive of research in numerous areas (linguistics, psycholinguistics, developmental psychology, second language research, speech pathology etc. etc.) over the last 50 years. These successes far outstrip anything that non-nativist learning theorists have able to achieve even in their wildest dreams, and support a powerful inference to the best explanation in the Chomskyan's favor.

**2.2.1(g) Who has the burden of proof?**

As seen above (§2.2.1(d)), however, the strength of the Chomskyan's ability to explain the phenomena of language acquisition has been questioned, and with it, implicitly, the strength of her inference to the best explanation. In addition, there is a general debate within the philosophy of science as to the soundness of inferences to the best explanation: does an explanation's being the best available give any additional reason (over and above its ability to account for the phenomena within its domain) to suppose it true? [Link to Encyclopedia Article 'Abduction' by Peter Achinstein for more on this topic.]

In the linguistic case, what sometimes seems to underpin people's positions on such issues is differing intuitions as to who has the burden of proof in this debate. Empiricists or non-nativists contend that Chomskyans have not presented enough data (or considered enough alternative hypotheses) to establish their case. Chomskyans reply that they have done more than enough, and that the onus is on their critics either to produce data disconfirming their view or to produce a testable alternative to it.

That such burden-shifting is endemic to discussions of linguistic nativism (the exchange in Ritter 2002 is illustrative) suggests to me that neither side in this debate has as yet fulfilled its obligations. Empiricists about language acquisition have ably identified a number of points of weakness in the Chomskyan case, but have only just begun to take on the demanding task of developing develop non-nativist learning theories, whether for language or anything much else. Nativists have rested content with hypotheses about language acquisition and innate knowledge that are based on plausible-seeming but largely unsubstantiated claims about what the *pld* contain, and about what children do and do not know and say.

It is unclear how to settle such arguments. While some may disagree (especially some Chomskyans), it seems that much work still needs to be done to understand how children learn language — and not just in the sense of working out the details of which parameters get set when, but in the sense of reconceiving both what linguistic competence consists in, and how it is acquired. In psychology, a new, non-nativist paradigm for thinking about language and learning has begun to emerge over the last 10 or so years, thanks to the work of researchers like Elizabeth Bates, Jeffrey Elman, Patricia Kuhl, Michael Tomasello and others. The reader is referred to Elman *et al*. 1996, Tomasello 2003 and §3 below for an entrée into this way of thinking.

For now, considerations of space demand a return to our topic, viz., linguistic nativism, rather than further discussion of alternatives to it.

## 2.3 The Argument from the 'Unlearning Problem'

We saw in the previous section that in order to support the view that all of UG is innately known, nativists about language need to hold not just that the data for language learning is impoverished in a few isolated instances, but that it's impoverished across the board. That is, in order to support the view that the innate contribution to language acquisition is something as rich and detailed as knowledge of Universal Grammar, nativists must hold that the inputs to language acquisition are defective in many and widespread cases. (After all, if the inputs were degenerate only in a few isolated instances, such as those discussed above, the learning problem could be solved simply by positing innate knowledge of a few relevant linguistic hints, rather than all of UG.)

Pullum and Scholz (2002:13) helpfully survey a number of ways in which nativists have made this point, including:

i. Finiteness: the *pld* (primary linguistic data) are finite, whereas languages contain infinitely many sentences.
ii. Underdetermination: the *pld* are always compatible with infinitely many grammatical hypotheses.
iii. Degeneracy: the *pld* contain ungrammatical and incomplete sentences.
iv. Idiosyncrasy: different children learning the same language are exposed to different samples of sentences.
v. Positivity: the *pld* contain only positive instances (what *is* a sentence of the language to be learned, a.k.a. the 'target language').
vi. No Feedback: children are not told or rewarded when they get things right, and are not corrected when they make mistakes.

In this section, I will set aside features (i) and (ii) as being characteristic of *any* empirical domain: the data are *always* finite, and they *always* underdetermine one's theory. No doubt it's an important problem for epistemologists and philosophers of science to explain how general theories can nonetheless be confirmed and believed. No doubt, too, it's an important problem for psychologists to explain the mechanisms by which individuals acquire general knowledge about the world on the basis of their experience. But underdetermination and the finiteness of the data are everyone's problem: if these features of the language learning situation per se supported nativism, then we should accept that all learning, in every domain, requires inborn domain-specific knowledge. But while it's not impossible that everything we know that goes beyond the data is a result of our having domain-specific innate knowledge, this view is so implausible as to warrant no further discussion here.

I also set aside features (iii) and (iv). For one thing, it is unclear exactly how degenerate the *pld* are; according to one early estimate, an impressive 99.7% of utterances of mothers to their children are grammatically impeccable (Newport, Gleitman and Gleitman 1977). And even if the data are messier than this figure suggests, it is not unreasonable to suppose that the vast weight of grammatically well-formed utterances would easily swamp any residual noise. As to the idiosyncrasy of different children's data sets, this is not so much a matter of stimulus poverty as stimulus difference. As such, idiosyncrasy becomes a problem for a non-nativist only on the assumption that different children's states of linguistic knowledge differ from one another less than one would expect given the differences in their experiences. As far as I know, no serious case for this last claim has ever been made.[12]

In this section, we will focus on features (v) and (vi) of the *pld*. For it is consideration of the positivity of the data set, and the lack of feedback available to children, that has given rise to what I am calling the 'Unlearning Problem,' otherwise known (somewhat misleadingly) as the 'Logical Problem of Language Acquisition.' (For statements of the argument, see, e.g., Baker 1979; Lasnik; 1989:89-90; Pinker 1989.)
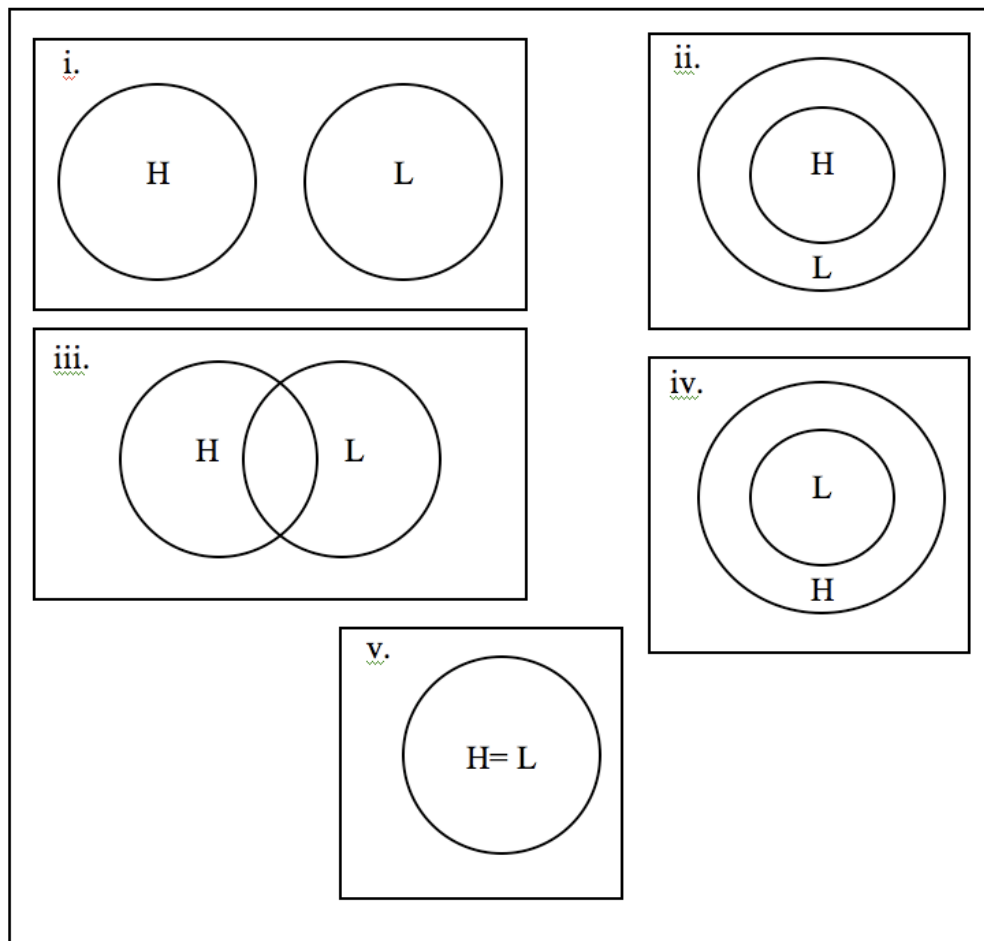
Figure 2. Five possible relations between the language generated by hypothesis (H) and the target grammar (L)

Take a child learning the grammar of her language, *L*. Figure 2 represents the 5 possible relations that might obtain between the language generated by her current hypothesis, *H*, and that generated by the target grammar, *L*. (v) represents the end point of the learning process: the learner has figured out the correct grammar for her language. A learner in situation (i), (ii) or (iii) is in good shape, for she can easily use the *pld* as a basis for correcting her hypothesis as follows: whenever she encounters a sentence in the data (i.e., a sentence of *L*) that is not generated by *H*, she has to 'expand' her hypothesis so that it generates that sentence. In this way, H will keep moving, as desired, towards *L*. However, suppose that the learner finds herself in situation (iv), where her hypothesis generates all of the target language, *L*, and more besides. (Children frequently find themselves in this position, for example, they invariably go through a phase in which they overgeneralize regular past tense verb endings to irregular verbs; their grammars generate the incorrect *\*I breaked it* as well as the correct *I broke it*.) There, she is in deep trouble, for she cannot use the *pld* to discover her error. Every sentence of *L*, after all, is already a sentence of *H*. In order to 'shrink' her hypothesis — to 'unlearn' the rules that generate *\*I breaked it* — she needs to know which sentences of *H* are not sentences of *L* — she needs to figure out that *\*I breaked it* is not a sentence of English. But — and this is the problem — this kind of evidence, often called 'negative evidence,' is held to be unavailable to language learners.

For as we have seen, the *pld* is mostly just a sample of sentences, of positive instances of the target language. It contains little, if any, information about strings of words that are not sentences. For instance, children aren't given lists of ungrammatical strings. Nor are they typically corrected when they make mistakes. And nor can they simply assume that strings that haven't made their way into the sample are ungrammatical: there are infinitely many sentences that are absent from the data for the simple reason that no-one's had occasion to say them yet.

In sum: a child who is in situation (iv) — a child whose grammar 'overgenerates' — would need negative evidence in order to recover from her error. Negative evidence, however, does not appear to exist. Since children do manage to learn languages, they must never get themselves into situation (iv): they must never need to 'unlearn' any grammatical rules. There are two ways they could do this. One would be never to generalize beyond the data at all. But clearly, children do generalize, else they'd never succeed in learning a language. The other would be if there were something that ensured that *when* they generalize beyond the data, they don't *overgeneralize*, something, that is, that ensures that children don't make errors that they could only correct on the basis of negative evidence. According to the linguistic nativist, this something is innate knowledge of UG.

## 2.3.1 Criticisms of the 'Unlearning' Argument

### 2.3.1 (a) What is lacking? Negative data vs. Negative Evidence

First, let's make a distinction between:

> *Negative Data*: explicit information that a given string of words is not a sentence of the target language. (E.g., "No, that's not how you say it," or "It's *I broke it* not *I breaked it*," or "That string of words is ungrammatical," etc.)

and

> *Negative Evidence*: information that would enable a learner to tell that a given hypothesis is (very likely to be) incorrect. (See below for examples.)

Second, let's abandon the idea, which reappears in many presentations of the Argument from the Unlearning Problem; that learners' hypotheses must be explicitly falsified in the data in order to be rejected. Let's suppose instead that learners proceed more like actual scientists do — provisionally abandoning theories due to lack of confirmation, making theoretical inferences to link data with theories, employing statistical information, and making defeasible, probabilistic (rather than decisive, all-or-nothing) judgments as to the truth or falsity of their theories.[13]

Intuitively, viewing the learner as employing more stochastic and probabilistic inductive techniques enables one to see how the unlearning problem might have been overblown. What the argument claims, rightly, is that *negative data* near enough do not exist in the *pld*. However, what learners need in order to recover from overgeneralizations, is not negative data *per se*, but *negative evidence*, and arguably, the *pld* do contain significant amounts of that. For example:

1. *Failures of understanding or communication*: others' failures to understand children's linguistic productions (evidenced either by requests for repetition or by communicative failure) are evidence to the learner that there is something wrong with the rule(s) she was using to generate her utterance. This evidence is not decisive (maybe Granny just couldn't hear her properly), but it is negative evidence nonetheless.

2. *Non-occurrence of structural types as negative evidence*: Suppose that a child's grammar predicted that a certain string is part of the target language. Suppose further that that string never appears in the data, even when the context seems appropriate. Proponents of the unlearning problem say that non-occurrence cannot constitute negative evidence — maybe Dad simply always chooses to say *The girl who is in the jumping castle is Kayley's daughter, isn't she?* rather than the auxiliary-fronted version, *Is the girl who is in the jumping castle Kayley's daughter?* If so, it would be a mistake for the child to conclude on the basis of this information that the latter string is ungrammatical.

   But suppose that the child is predicting not strings of words, *simpliciter*, but rather strings of words *under a certain syntactic description* (or, perhaps more plausibly, quasi-syntactic description — the categories employed need not be the same as those employed in adult grammars).[14] This would enable her to make

much better use of non-occurrence as negative evidence. For non-occurring strings will divide into two broad kinds: those whose structures *have* been encountered before in the data, and those whose structures have *not* been heard before. In the former case, the child has positive evidence that strings of that kind are grammatical, evidence that would enable her to suppose that the non-occurrence of that particular string was just an accident. (E.g., she could reason that since she's heard *Is that mess that is on the floor in there yours?* many times, and since that string has the same basic structure as *Is that girl that's in the jumping castle Kayley's daughter?*, the latter string is probably OK even though Dad chose not to say it.)

In the case in which the relevant form has *never* been encountered before in the data, however, the child is better off: the fact that she has never heard any utterance with the structure of *\*Is that girl who in the jumping castle is Kayley's daughter* or *\*Is that mess that that on the floor in there is yours?* is evidence that strings of that type are not sentences. Again, the evidence is not decisive, and the child should be prepared to revise her grammar should strings of that kind start appearing. Nonetheless, the non-occurrence of a string, suitably interpreted in the light of other linguistic information, can constitute negative evidence and provide learners with reason to reject overgeneral grammars.

3. *Positive Evidence as Negative Evidence*. Relatedly, learners can also exploit positive evidence as to which strings occur in the *pld* as a source of negative evidence — again in a tentative and revisable way.[15] Suppose that the child's grammar generated two strings as appropriate in a given kind of context, but that only one sort of string was ever produced by those around her. The fact that only strings of the first kind occur is in this case negative evidence — defeasible, to be sure, but negative evidence nonetheless.

In fact, the use of positive evidence to disconfirm hypotheses is endemic to science. For instance, Millikan used positive evidence to disconfirm the theory that electrical charge is a quantity that varies continuously. In his famous 'Oil Drop' experiment, he found that the amount of charge possessed by a charged oil drop was always a whole-number multiple of $-(1.6 \times 10\text{-}19)$C. The finding that all observed charges were 'quantized' in this manner disconfirmed the competing 'continuous charges' hypothesis in the same way that positive evidence can disconfirm grammatical hypotheses.[16]

4. *Feedback* The Argument from the 'Unlearning Problem' also points to the lack of feedback provided to children learning language. In a famous study often cited by proponents of the argument, Brown and Hanlon 1970 (see also Brown, 1973 and Brown, Cazden, and Bellugi 1969) found no overt disapproval by mothers of the syntactic errors of their children, and moreover found that caregivers had no trouble understanding their charges' ill-formed utterances. Only semantic errors were occasionally corrected; grammatical mistakes went unremarked.

However, more recent findings have uncovered evidence indicating that failures of understanding occur with some regularity, and that there is a wealth of feedback about correct usage in the language-learning environment. For example:

- Hirsh-Pasek, Trieman and Schneiderman (1984) studied interactions between 2 year olds and their parents, and discovered that caregivers repeated and corrected 20.8% of flawed sentences, whereas they only repeated (without correction) 12.0% of well-formed utterances.
- Demetras, Post and Snow (1986) found that in general, only well-formed sentences were repeated verbatim by parents, and that ill-formed sentences were not repeated verbatim, but were rather followed by clarification questions ("What?" — indicating a lack of understanding) or expansions and/or recasts, correcting the error.
- Bohannon and Stanowicz (1988) found that 34% of syntactic and 35% of phonological errors received some form of differential feedback (e.g., repetitions with corrections or explicit rejection of the utterance); that more than 90% of parents' exact repetitions follow well-formed sentences; and that more than 70% of recasts and expansions follow ill-formed utterances.
- Chouinard and Clark's (2003) longitudinal study of five children learning language found that parents reformulate erroneous utterances more often than correct utterances, that they respond equally often to all

error types (phonological, lexical, syntactic, semantic), and that they correct younger children, who make more errors, more frequently.

- Perhaps most tellingly, Moerk (1991) performed a reanalysis of Brown's "Eve" transcripts (among those on which the 1970 'no feedback claim was based) and found many instances in which Eve's semantic and syntactic errors were explicitly corrected, including: her use of noun labels; VPs (tense, modality, auxiliaries); determiners and prepositions; word order (these last sorts of error were rare, but were invariably corrected).
- Bohannon, MacWhinney and Snow (1990) review other results in this vein, as well as responding to nativist criticisms of these findings and their bearing on the unlearning problem.

### 2.3.1 (b) Children Can and Do Learn from 'Noisy' Data and Exploit Statistical Regularities

Chomsky has recognized the existence of such 'indirect' negative data in the *pld*. However, he concluded that they were too few and ambiguous to be of aid to the language learner. The sorts of findings reported above seem to show that negative evidence is pervasive in the *pld*. But can children learn from these sorts of statistical regularities?

Standard formulations[17] of the 'Unlearning Problem', assume that they cannot: the view seems to be that learning can only take place under idealized conditions where the world supplies unambiguous evidence pro or con the language learner's grammatical theories. Given such a conception of the learner, none of the examples of feedback just discussed will seem relevant to the problem. For only a learner employing fairly sophisticated data-analysis techniques and a confirmation measure that is sensitive to small changes in probabilities would be able to exploit the sorts of regularities in the linguistic environment that we have just discussed. However, there is increasing evidence that children are in fact remarkably sensitive to subtle feedback, in both linguistic and non-linguistic domains. For instance:

- Bohannon and Stanowicz (1988) found that children pay particular attention to parental corrections of their mistakes: they imitate 25.6% of adult expansions (saying the same thing as the child, but giving more detail) and recasts (repetitions of the child's utterance, correcting errors), whereas they only imitate 3.6% of exact or verbatim repetitions by parents of the child's utterance.
- Relatedly, Farrer (1990 1992) found that children were more likely to repeat a given morpheme if it were part of an adult recast of one of the child's own sentences, than if it were part of a a non-repetitious adult utterance (e.g., a change of subject or a continuation of the conversation). She also found that children's repetition of adult utterances facilitated the child's acquisition of various grammatical morphemes.
- Morgan and Travis (1989) and Morgan *et al.* (1995) dispute the long-term efficacy of such corrective feedback; Bohannonn *et al.* 1996 respond.
- However, in both longitudinal studies of children in natural environments (Chouinard and Clark 2003) and in experimental studies (Saxton *et al.* 1998, and Saxton 1997, Saxton, Backley and Galloway, 2003) the long-term efficacy of feedback has been demonstrated.

In addition, it is becoming increasingly clear that babies, children, adults and many other mammals are highly sensitive not just to feedback, but to other non-obvious statistical regularities in their experience. For example:

- Saffran, Aslin and Newport (1996) found that 8 month old babies were able to learn where the word boundaries in an artificial language occurred after a mere 2 minutes' exposure to a stream of artificial speech. The stream consisted of 3-syllable nonsense words (*bidaku*, *padoti*, *golabu*) repeated continuously for 2 minutes (*bidakupadotigolabubidakugolabi* …etc.). The stream was constructed so that the 'transitional probability' of two sounds X#Y was equal to 1 when the sounds formed part of a word, and equal to 1/3 when the sounds spanned a word boundary. In two minutes, the infants had learned to discriminate the 'words' (like *bidaku)* from the 'non-words' (e.g., *kupado*). See also Chambers, Onishi and Fisher 2003.
- Other studies have expanded upon these results, indicating that children and babies are sensitive to patterns in a wide range of verbal cues, such as linguistic rhythm (Nazzi and Ramus 2003); prosodic stress (Thiessen and Saffran 2003) and voicing and syllabic structure (Saffran and Thiessen 2003).

- Moreover, there is increasingly persuasive evidence that statistical or 'distributional' information may be used not just for the extraction of word boundaries, but — contrary to an old argument of Chomsky's — to limn higher levels of syntactic structure as well — see, e.g., Redington and Chater, 1998; Pena *et al.* 2002; Mintz 2002; Saffran 2002; Saffran and Wilson 2003; Newport and Aslin 2004. Chater and Manning 2006 provide a survey.)
- Finally, and forestalling any response along the lines that what we are seeing here is just the nativist's 'Language Acquisition Device' in action, a number of studies have shown that similar mechanisms appear to be at work in learning in non-linguistic domains (Saffran 2002, studied learning of non-linguistic sounds and shapes); in adults (Pena *et al.* 2002); and in other animals, such as cotton-top tamarin monkeys (Hauser 2001; Hauser, Weiss and Marcus, 2002).

Taken together, these kinds of results raise the possibility that some of the foundational learning mechanisms involved in language acquisition are not language specific. If it turns out that babies employ the sorts of distributional analysis studied by Saffran, Redington and Chater, Pena, and Mintz not only in learning artificial languages, but also in learning natural languages, then that is evidence against linguistic nativism. For this type of learning is employed by humans and other animals in other contexts as well: whatever is involved in language learning — be it innate or not — is not language-specific.

### 2.3.1 (c) The Generality of the Argument

The previous objections to the Unlearning Problem Argument made the points, first, that negative evidence does exist in the *pld* (in the form of regularities both in others' language use and in how others react to children's own productions), and second, that children (and other animals) seem very good at exploiting this kind of information for the purposes of learning about their world. This would seem to be rather a good thing, given that there is reason to think that learners *must* be able to learn in domains where explicit negative data do not exist, and in the absence of specialized innate knowledge of those domains. For the unlearning problem is a problem for learning from experience quite generally. That is, there are many domains in which learners lack explicit evidence as to what things are *not*: trees are not cars, Irish stews are not curries, birds are not fish and MacDonald's is not a branch of the CIA. No-one ever told you any of these things, but it's crazy to think that you now know them because you possess analogs to the 'Language Acquisition Device' for each of these domains. Clearly, in at least some areas, people are able to learn an awful lot on the basis of largely positive data, and while this of course does nothing to show that language is one of those areas, it does indicate that the Unlearning problem argument by itself is no argument for linguistic nativism at all, let alone for the Chomskyan (UG-based) version of that position.

# 3. Other Research Bearing on the Innateness of Language: New Problems for the Nativist?

In this section, I will mention some other avenues of research that have been argued to have a bearing on the innateness of language. My goal is not to give an exhaustive survey of these matters, but rather to provide the interested reader with a way into the relevant literatures. Still, I will try to give enough details so as to make a case that current empirical findings, together with the flaws identified in §§1 and 2 in the positive arguments for linguistic nativism, tend to militate against that position.

## 3.1 Linguistic Universals

Chomsky and others (e.g., Chomsky 1988:46-7; Pinker 1994:237-8) have pointed to the existence of 'linguistic universals' as supporting the idea that language is the product of a distinct faculty of mind. Universals are features thought to be common to all natural languages, such as the existence of constraints on the movement of elements during a derivation or, less controversially, the existence of a syntactic distinction between nouns and verbs. But not only is the existence of true universals a contested matter (see e.g., Maratsos 1989:111), it is unclear what the correct explanation of them — assuming they exist — is would be.

One explanation is certainly the Chomskyan one that they are consequences of speakers' innate knowledge of UG. Another is that they derive from other, non-linguistically-specific features of cognition, such as memory or processing constraints (e.g., Berwick and Weinberg 1983 trace certain constraints on movement to limitations on parsing imposed by the structure of human memory). Yet another is that they derive from universal demands of the communication situation (e.g., Sapir 1921, argued that the distinction between nouns and verbs arises from the fact that language is used to communicate propositions, hence needs a way to bring an object subject to mind and a way to say something about it). Finally, as Putnam 1971 speculated, universals might be relics of an ancestral Ur-language from which all other languages evolved. This last hypothesis has generally been rejected as lacking in empirical support. However, recent findings in genetics and historical linguistics are converging to suggest that all human populations evolved from a small group migrating from Africa in the fairly recent past, and that all human languages have probably evolved from the language spoken by that group. (Cavalli-Sforza 1997.)

The Ur-language hypothesis is not, of course, inconsistent with linguistic nativism. However, if true, it does weaken any *argument* from the existence of universals to the innateness of linguistic knowledge. For if languages have a common ancestor, then it is possible to explain universals — even ones that seem strange from a functional point of view — as being the result of our ancestors' having adopted a certain solution to a linguistic coordination problem. Like driving on the right side of the road, a solution once established may become entrenched, because the benefits of everyone's conforming to the same rule outweigh the costs of changing to a different rule, and this may be so even if the new rule were in some sense more 'reasonable.' Thus, even arbitrary or odd features of language can be explained historically, without positing either compelling functional considerations or inborn linguistic constraints.[18]

If, by contrast, language emerged independently in a number of areas, the existence of universals would be a strong argument for nativism, For in that case, it would be implausible to maintain that each ancestral group 'just happened' to select the same solutions to the various coordination problems they encountered. More plausible would be the supposition that the different groups' choice of the rule was driven by something *internal to speakers*, such as, perhaps, an innate representation of UG. In short: if languages have a common ancestor, then common descent from originally arbitrary linguistic conventions is a possible explanation of linguistic universals, including the 'odd' or 'arbitrary' ones that don't seem to have any real functional significance. If they don't, then such universals seemingly could only be explained in terms of features internal to speakers.
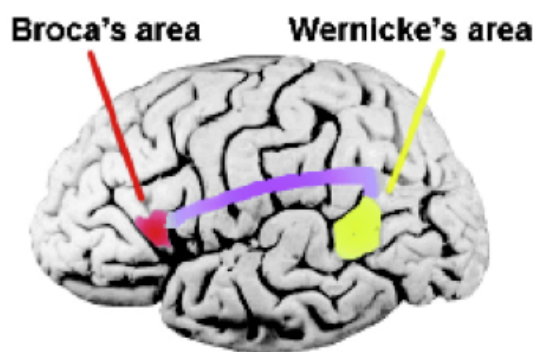
## 3.2 Language Localization



Figure 3. Broca's area and Wernicke's area

Beginning with the work of Broca and Wernicke in the 19th century, a popular view has been that language is localized to certain areas of the brain (see Fig. 3), almost always the left hemisphere,[19] and that it is subject to characteristic patterns of breakdown, called 'aphasias.' (See Saffran 2000 for a survey of the various aphasias.) For example, Broca's area is strongly implicated in speech production, and damage to this area can result in a characteristic inability ('Broca's aphasia' or 'agrammatism') to produce fluent speech, especially complex grammatical structures and grammatical morphemes. The fact that syntax can apparently be selectively interfered with by lesions to Broca's area has been taken by some to indicate that grammatical knowledge is localized to

that area, and this in turn has been taken to show support the view that there is a special biological inborn basis for that knowledge. (Lenneberg 1964, 1967 is the original proponent of this argument, which is echoed in more recent discussions, such as Pinker 1994:297-314.)

It is unclear, however, why this inference should seem compelling. First, as Elman *et al.* 1996 argue, neural localization of function can occur as a result of virtually any developmental trajectory: the localization of some function bears not at all on its innateness.

Secondly, it is now known that neural localization for language is very much a relative, rather than an all-or-nothing matter (Dronkers *et al.* 2000, Dick *et al.* 2001, Martin 2003). Not only is language processing widely distributed over the brain (see Fig. 4), but traditionally language-specific areas of cortex are implicated in a variety of non-linguistic tasks as well. Broca's area, for instance, 'lights up' on MEG scans (magnetoencephalography, a method for measuring changes in the magnetic properties of the brain due to electrical activity) when subjects hear a discordant musical sequence in much the same way as it does when they hear an ungrammatical utterance. (Maess *et al.* 2001; a special issue of *Nature Neuroscience*, 6(7), July 2003, explores the implications of this finding.)

Finally, recent studies of cortical plasticity have shown that even the most plausible candidates for innate specification — such as the use of visual cortex for vision or the use of auditory cortex for hearing — exhibit high degrees of experience-dependent plasticity. For example, in congenitally blind subjects, the areas of the brain normally used for seeing are taken over for the processing of Braille (Sadato *et al.* 1996; Hamilton and Pascual-Leone, 1998) and even in those with late-onset blindness, significant 'rewiring' of visual cortex for other perceptual tasks is apparent (Kujala *et al.* 1997). Likewise, in the congenitally deaf, auditory cortex is used for the processing of sign language (Nishimura *et al.* 1999, von Melchner, Pallas and Sur 2000). (See Shimojo and Shams 2001, for a review.)
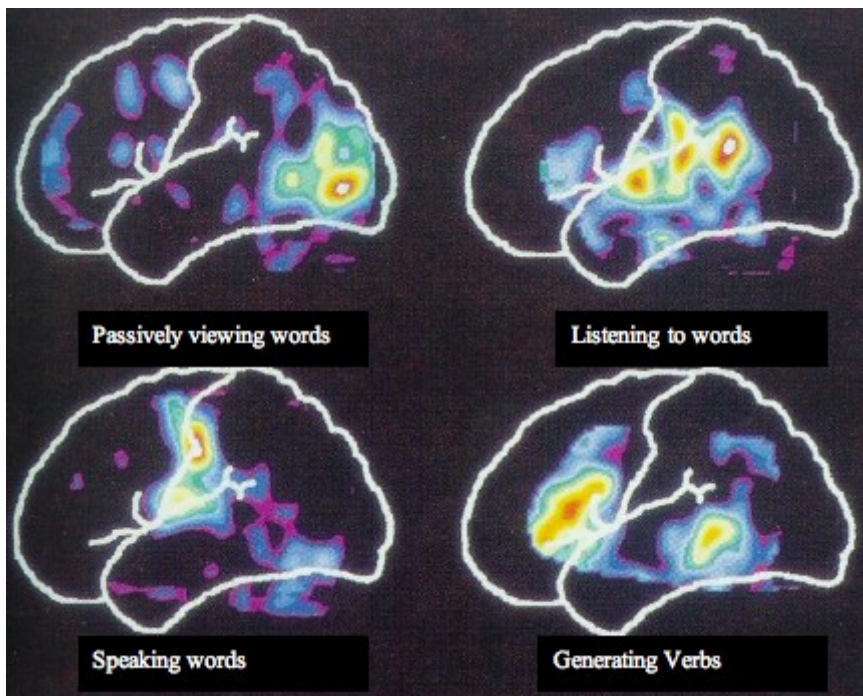


Figure 4. Pet scan showing brain regions involved in various language tasks. From Posner and Raichle (1997, 15). Used by permission of M. Raichle.

As Marcus (2004:40-45) points out in response to Elman *et al.* 1996, the ability of the brain to 'rewire' itself under exceptional circumstances is consistent with its having been 'prewired,' or set up, differently by the genes. However, these sorts of data indicate that complex functions, such as are involved in processing sign language, can be carried out in areas of brain that are 'prewired' (if they are) to do something quite different. This suggests that these abilities require little in the way of task-specific pre-wiring, and are learned largely on the basis of

experience (together with whatever sort of 'prewiring' is supplied for the cortex as a whole). That is, if sign language processing tasks can be carried out by areas of cortex that are presumably innately predisposed (if they are) to do auditory processing, then the former competence must be being learned in the absence of inborn constraints or knowledge that are specific to that task. Of course, these are pathological cases, and it is unclear whether the subjects in these experiments had any special training in order that their brains were 'rewired' in these ways. Nonetheless, examples like these provide an existence proof of the brain's ability to acquire complex processing capacities — indeed, processing capacities relevant to language — in the complete absence of inborn, domain-specific information. As such, they raise the possibility that other aspects of language processing are similarly acquired in the absence of task-specific constraints.

In sum, the neuroscientific evidence currently available provides no support for linguistic nativism. The suggestion that localization of function is indicative of a substantial degree of innate prespecification is no longer tenable: localization can arise in many different ways. In addition, linguistic functions do not seem to be particularly localized: language use and understanding are complex tasks, involving many different brain areas — areas that are in at least some cases implicated also in other tasks. It is hard to see how to reconcile these facts with the Chomskyan postulation of a monolithic 'language organ,' the development or 'growth' of which is controlled largely by the genes. Finally, the fact that complex functions can be learned and carried out by areas of brain that are innately 'prewired' (if at all) to do quite different sorts of processing indicates that such competences can be and are acquired without any inborn, task-specific guidance. This is not, of course, to say that language is one of the competences that are acquired in this way. For all the current evidence shows, many areas of cortex in which language develops may indeed be 'prewired' for that task: linguistic nativism is still consistent with what is now known. It is, however, to suggest that although there may be other reasons to be a linguistic nativist, general considerations to do with brain organization or development as currently understood give no especial support to that position.

## 3.3 The Critical Period for Language Acquisition

Lenneberg (1964, 1967) also argued that although language acquisition is remarkably robust, in the sense that all normal (and many abnormal) children do it, it can occur unproblematically only during a 'critical period' — roughly, up to late childhood or early puberty. On analogy with other supposedly innately specified processes like imprinting or visual development, Lenneberg used the existence of a critical period as further evidence that language possesses a proprietary basis in biology.

In support of the critical period hypothesis about language, Lenneberg cited the facts (i) that retarded (e.g., Downs syndrome) children's language development stops around puberty; (ii) that whereas very young children are able to (re)learn language after aphasias produced by massive left-hemisphere trauma (including hemispherectomy), aphasias in older children and adults are typically not reversible; and (iii) that so-called 'wild children,' viz., those who grow up with no or little exposure to human language, exhibit severely compromised language skills. (Lenneberg, 1957:142-55; see Curtiss 1977 for the (in)famous case of Genie, a modern-day 'wild child' from suburban Los Angeles, who was unable to acquire any but the most rudimentary grammatical competence after a miserable and wordless childhood spent locked alone in a room, tied to her potty chair or bed.)

As further support for the critical period hypothesis, others have added the observation that although children are able to learn a second language rapidly and to native speaker fluency, adult learners of second languages typically are not: the capacity to learn a second language tapers off after puberty, no matter how much exposure to the language one has. (Newport 1990). Thus, it was speculated, the innate knowledge base for language learning (e.g., knowledge of UG) becomes unavailable for normal acquisition at puberty, and adult learners must rely on less efficient learning methods. (Johnson and Newport 1989.)

As a preliminary to discussing these arguments (many of which are presented in more detailed in Stromswold 2000) it is worth distinguishing two notions that often get conflated under the name 'critical period':

*Critical Period*: a time during development which is *literally* critical; the relevant competence either cannot develop or will be permanently lost unless certain inputs are received during that period.

*Sensitive Period*: a time during development in which a competence is acquired 'normally,' or 'easily,' or 'naturally.' The competence can be acquired outside the sensitive period, but perhaps less easily and naturally, and or perhaps with less ultimate success.

The classic example of a critical period is due to the Nobel prize-winning work of Hubel and Wiesel. By suturing shut one of a kitten's eyes at various stages of development and for various periods of time, Hubel and Wiesel (1970) showed that certain cortical and thalamic areas supporting binocular vision (specifically, ocular dominance columns[20] and cells in the lateral geniculate body) will not develop normally unless kittens receive patterned visual stimulation during the 4th to 12th weeks of life. They found that while the damage was sometimes reversible to some extent, depending on the exact duration and timing of the occlusion, occlusion for the entire first three months of life produced irreversible blindness in the deprived eye.[21]

Language, however, is not like this. As we will see, there is little evidence for a *critical period* for language acquisition, although there is considerable evidence that there is a *sensitive period* during which language is acquired more easily. The implications of this for claims about the innateness of language will be addressed in §3.3.4.

### 3.3.1 Language recovery after trauma

Lenneberg cited the superior ability of children to (re)learn language after left brain injury in support of the critical period hypothesis. But while there clearly is a difference between the abilities of young children, on the one hand, and older children and adults, on the other, to recover from left brain insults, the contrast in recovery course and outcome is not as stark as is often supposed.

First, older children — even those who have not succeeded in learning language previously — can substantially recover from left hemisphere trauma occurring well after the supposed closure of the 'sensitive' or 'critical' period; in effect, they learn language from scratch as adolescents. Vargha-Khadem *et al*. 1997, for instance, report the case of Alex, who failed to speak at all during childhood and whose receptive language was at age 3-4 level at age 9. After his left cortex was removed at age 9, Alex suddenly began to learn language with gusto, and by age 15, his skills were those of an 8-10 year old.

Secondly, most adults suffering infarcts in the left hemisphere language areas do in fact recover at least some degree of language competence and many recover substantially normal competence, especially with treatment (Holland *et al*. 1996). This is thought to be due both to the regeneration of damaged speech areas and to compensatory development in other areas, particularly in the right hemisphere (Karbe *et al*. 1998). Similar processes seem to be at work in young children with left hemisphere damage. Muller *et al*. 1999, for instance, document significant relearning of language, together with increased right-hemisphere involvement in language tasks, after left-hemisphere lesions in both children (<10) years) and adults (>20 years).

Finally, not even very young children are guaranteed to recover language after serious insults, whether to the left or right hemisphere. As Bates and Roe (2001) argue in their survey of the childhood aphasia literature, outcomes differ wildly from case to case, and the reported studies exhibit numerous methodological confounds (e.g., inability to localize the lesion or to know its cause, different measures of linguistic competence, different time frames for testing, statistical irregularities, and failure to control for other factors known to affect language such as seizure history) that cast doubt on the degree of empirical support possessed by Lenneberg's claim in this instance.

### 3.3.2 'Wild children'

It has long been recognized that interpretation of the 'wild child' literature — helpfully surveyed in Skuse 1993 — is confounded by the fortunate rarity of these 'natural experiments,' the generally poor reporting of them, and the other environmental factors (abuse, malnutrition, neglect, etc.) that often go along with extreme linguistic deprivation. However, in work pioneered by Goldin Meadow and colleagues (e.g., Goldin Meadow and Mylander 1983, 1990), a new population of individuals, who are linguistically but not otherwise deprived, has begun to be studied. Deaf but otherwise normal children of hearing parents who are neither educated in sign language nor sent to special schools for the deaf do not acquire language, although they usually develop their own rudimentary signing systems, called 'homesign,' to use with their families. Studies of what happens to such children after they are exposed to natural languages (signed or verbal) at various ages promise to offer new insights into the critical and sensitive period hypotheses.

At this time, however, there are still very few case reports in the literature, and the data so far obtained in these studies are equivocal with respect to the sensitive and critical period hypotheses. Some adolescents do seem to be able to acquire language despite early linguistic deprivation, and others do not. It is unclear what the explanation of these different outcomes is, but one important factor appears to be whether the new language is a signed language (e.g., ASL) or a spoken language. Perhaps because their childhood perceptual deficits prevented normal auditory and articulatory development, deaf children whose hearing is restored later in life do not seem to be able to acquire much in the way of spoken language. (Grimshaw *et al*. 1998.)

### 3.3.3 Second language acquisition in children and adults

The issue of second language acquisition ("SLA") has been argued to bear on the innateness of language by supporting a critical (or sensitive) period hypothesis. For instance, Johnson and Newport (1989) found that among immigrants arriving in the U.S. before puberty, English performance as adults was better the earlier in life they arrived, but that there were no effects of arrival age on language performance for those arriving after puberty. The fact that the amount of exposure to the second language mattered for speakers if it occurred *before* puberty but not *after,* was taken to confirm the critical period hypothesis.

However, these results have failed to be replicated (Birdsong and Molis 2001) and while it still has its supporters, the 'critical period' hypothesis regarding second language acquisition is increasingly being criticized (Hakuta, Bialystok and Wiley 2003; Nikolov and Djugunovich 2006). Newer studies have have argued, for instance, that the degree of proficiency in a second language correlates better with, such factors as the learner's level of educational attainment in that language, her length of residence in the new country,) and the grammatical similarities between the first and second languages, and/or length of residence in the new country. (Flege, Yeni-Komshian and Liu 1999; Bialystok, 1997)[22]

The fact that many adults and older children can learn both first and second languages to a high degree of proficiency makes clear that unlike the kitten visual system studied by Hubel and Wiesel, the language acquisition system in humans is not subject to a critical period in the strict sense. This finding is consistent with the emerging view that the cortex remains highly plastic throughout life, and that contrary to received wisdom, even old dogs can be quite good at learning new tricks. (See Buonomano and Merzenich 1998; Cowen and Gavazzi 1998; Quartz and Sejnowski 1997; and Stiles 2000.) It is also consistent with the idea, which seems more plausible than the critical period hypothesis, that there is a *sensitive* period for language acquisition — a time, from roughly birth age 1 to age 6 or 7, in which language is acquired most easily and naturally, and when a native-like outcome is virtually guaranteed. (Cf. Mayberry and Eichen 1991.) The implications of this conclusion for linguistic nativism are examined in the next section.

### 3.3.4 Sensitive periods and innateness: phonological learning

What does the existence of a sensitive period for language mastery tell us about the innateness of language? In this section, we will look at a case, namely phonological learning, in which the existence of a sensitive period has received much press, and in which the inference from sensitivity to the existence of language-specific innate

information has been made explicitly (see Eimas 1975). One can argue that even in this case, the inference to linguistic nativism is weak.

Much rarer than mastery of second language morphology and syntax is attainment of a native-like accent, something that first language learners acquire automatically in childhood.[23] A child's ability to perceive language-specific sounds begins in utero, as demonstrated, for instance, by newborns' preference for the sounds of their mother's voice and their parents' language, and by their ability to discriminate prose passages that they have heard during the final trimester from novel passages. In the first few months of life, babies reliably discriminate many different natural language phonemes, whether or not they occur in what is soon to become their language. By ages 6 months to 1 year, however, this sensitivity to unheard phonemes largely disappears, and by age 1, children tend to make only the phonological distinctions made in the language(s) they hear around them. For example, Japanese children lose the ability to discriminate English /r/ and /l/ (Kuhl *et al.*, 1997b). As adults, people continue to be unable to perceive some phonetic contrasts not marked by their language, and many fail to learn how to produce even those second language sounds which they can distinguish.[24] For instance, many English speakers of French have great difficulty in producing the French /y/ (as in *tu*) and back-of-the-throat /r/.

Thus, in the case of phonological learning, there does seem to be an inborn predisposition to segment vocal sounds into language-relevant units, or phonemes.[25] However, there is also evidence that learning plays a role in shaping phonological knowledge — and not just by 'pruning away' unwanted 'phonological representations,' as Eimas (1975) hypothesized, but also by shaping the precise boundaries of adult phonemic categories. For example, caregivers reliably speak a special 'language' ("Motherese" or "Parentese") to young babies, raising pitch, shortening sentences, emphasizing stressed morphemes and word boundaries and — most relevant here — exaggerating the acoustical differences between certain crucial vowels (in English, /i/, /a/ and /u/) . This 'stretching' of the distance between vowels (demonstrated in Finnish and Russian as well as English by Kuhl et al. 1997a) facilitates the infant's representation of clearly distinguishable vowel prototypes. Kuhl 2000 argues that these prototypes subsequently function as 'magnets' around which subsequent linguistic experiences are organized, and form the set points of the language-specific phonological 'map' that emerges by the end of the first year.

If this is indeed how phonological learning works, it is clear that while experience clearly plays a role, the inborn contribution to that process is quite substantial. For discriminating phonemes — however those discriminations might be shaped by subsequent experience — is no simple matter. It involves what is called 'categorical perception, that is, the segmenting of a signal that varies continuously along a number of physical dimensions (e.g., voice onset time and formant frequency) into discrete categories, so that signals within the category are counted as the same, even though acoustically, they may differ from one another more than do two signals in different categories (see Fig. 5). (Harnad 1987 is a useful collection of work on categorical perception to the mid-1980s.)

But is this inborn contribution to phonological learning *language specific*, that is, does it support the conclusion that (this aspect of) language is innate? And to this question, the answer appears to be 'No.' First, the 'chunking' of continuously varying stimuli into discrete categories is a feature not just of speech perception, but of human perception generally. For instance, it has been demonstrated in the perception of non-linguistic sounds, like musical pitch, key and melody, and meaningless chirps and bleats (Pastore and Layer 1990). It has also been demonstrated in the processing of visual stimuli like faces (Beale and Keil 1995), facial expressions (Etcoff and Magee 1992; Kotsoni, de Haan and Johnson 2001); facial gender (Campanella, Chrysochoos and Bruyer 2001); and familiar physical objects (Newell and Bulthoff 2002). Secondly, it is known that other animals too perceive categorically. For instance, crickets segment consepecific songs in terms of frequency (Wyttenbach, May and Hoy 1996), swamp sparrows 'chunk' notes of differing durations (Nelson and Marler 1989), and rhesus monkeys can recognize melodies when transposed by one or two octaves, but not by 1.5 or 2.5 octaves, indicating a grasp of musical key (Wright *et al*. 2000). Finally, other species respond categorically to human speech! Chinchillas (Kuhl and Miller 1975) and cotton-top tamarins (Ramus *et al*. 2000) make similar phonological distinctions to those made by human infants.

Together, as Kuhl 1994, 2000 argues, these findings cast doubt on the language-specificity of the inborn perceptual and categorization capacities that form the basis of human phonological learning. For given the fact that human (and animal) perception quite generally is categorical, it is arguable that languages have evolved so as to exploit the perceptual distinctions that humans are able to make, rather than humans' having evolved the abilities to make just the distinctions that are made in human languages, as a view like Eimas' would suggest.
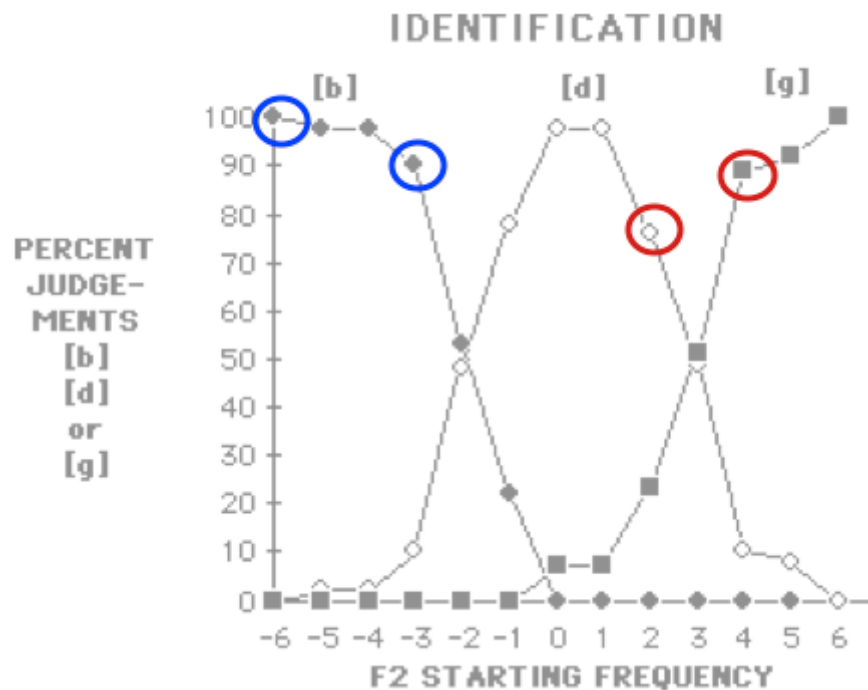


Figure 5. Note that the pair of sounds circled in blue differ in F2 starting frequency less than those circled in red, yet the former are both reliably counted as instances of the sounds /b/ whereas the latter are reliably classified as different sounds, /d/ and /g/. This pattern, together with the abrupt switch from one classification to another (e.g. /b/ to /g/), is characteristic of categorical perception.

The same may be true in non-phonological domains too. The notion that at least some of the capacities responsible for syntactic learning are non-language specific is suggested by analogous results about the non-species specificity of recursive rule learning and generalization — an ability that Chomsky has recently suggested forms the core of the human language faculty. (Hauser, Chomsky and Fitch 2002; see below, 3.4 for further discussion.) Other species, notably cotton top tamarins, seem capable of learning simple recursive rules (Hauser, Weiss, and Marcus 2002). In addition, Hauser and McDermott 2003 argue that musical and syntactic processing involve similar competences, which are again seen in other species. Together, these findings suggest that there are aspects of the human 'language faculty' that are neither task-specific nor species-specific. Instead, language learning and linguistic processing make use of abilities that predate language phylogenetically, and that are used in humans and in animals for other sorts of tasks. (See e.g., Hauser, Weiss, and Marcus 2002 for an account of recent work on rule learning by cotton top tamarins; see Hauser and McDermott 2003 for the suggestion that aspects of musical and syntactic processing involve similar competences, which are again seen in other species.) Rather than viewing the human mind as being innately specialized for language language learning, it seems at least as reasonable to think of languages as being specialized so as to be learnable and usable by the human mind; of this, more in §3.4 below.

## 3.4 Language Evolution

This brings us to the question of language evolution: if knowledge of language (say, of the principles of UG) really is inborn in the human language faculty, how did such inborn knowledge evolve? For many years, Chomsky himself refused to speculate about this matter, stating that "[e]volutionary theory…has little to say, as of now, about questions of this nature" (1988:167). Other theorists have not been so reticent, and a large

literature has grown up in which the selective advantages of having a language are adumbrated. It's good for communicating with, for instance, when trying to figure out what conspecifics are up to (Pinker and Bloom, 1990; Dunbar 1996). It's a mechanism of group cohesion, analogous to primate grooming (Dunbar 1996). It's a non-genetic mechanism of phenotypical plasticity, allowing organisms to adapt to their environment in non-evolutionary time (Brandon and Hornstein 1986; Sterelny 2003). It's a mechanism by which we can bend others to our will (Dawkins and Krebs 1979; Catania 1990), or make social contracts (Skyrms 1996). Language makes us smarter, perhaps by being internalized and functioning as a 'language of thought' (Bickerton 1995, 2000). And so on.

The ability to speak and understand a language no doubt provided and continues to provide us with many of these benefits. Consequently (and assuming that the costs were not too great — as patently they weren't), one can be sure that whatever it is about human beings that enables them to learn and use language would have been subjected to strong positive selection pressure once it began to emerge in our species.

But none of this speaks directly to the issue of linguistic nativism. The fact that Mother Nature would have favored individuals or groups possessing linguistic abilities tells us nothing about the *means* she chose to get the linguistic phenotype built. That is, it tells us nothing about the sorts of psychological mechanisms that were recruited to enable human beings to learn, and subsequently use, a natural language.

Nativism is, of course, one possibility. Natural selection might have built a specialized language faculty, containing inborn knowledge about language (e.g., knowledge of UG), which subsequently was selected for because it helped human children to acquire linguistic competence, and having linguistic competence enhanced our ancestors' fitness. A problem with this hypothesis, however, is that it is unclear how a language faculty containing innate representations of UG might have arisen in the human mind. One view is that the language faculty was built up piecemeal by natural selection. This approach underlies Pinker and Bloom's (1990) and Jackendoff's (1999) proposals as to the adaptive functions of various grammatical features and devices. Other nativists, however, reject the adaptationist framework. For instance, Berwick 1998, has argued that efforts to explain the piecemeal development of knowledge of linguistic universals in our species may be unnecessary in light of the new, Minimalist conception of syntax (see Chomsky 1995). On this view, all parametric constraints and rules of syntax are consequences of a fundamental syntactic process called Merge: once Merge was in place, Berwick argues, the rest of UG automatically followed. Chomsky, taking another tack, has suggested that language is a 'spandrel,' a byproduct of other non-linguistically directed selective processes, such as "the increase in brain size and complexity" (1982:23). And finally Bickerton 1998, on yet another tack, posits a massive saltative episode in which large chunks of syntax emerged all at once, although this posit is implicitly withdrawn in Calvin and Bickerton 2000.

The literature on language evolution is too large to survey in this article (but see Botha 2003 for an excellent overview and critique). Suffice it to note that as yet, no consensus has emerged as to how innate knowledge of UG might have evolved from whatever preadaptations existed in our ancestors. Of course, this is not in itself a problem for linguistic nativists: formulating and testing hypotheses about human cognitive evolution is a massively difficult enterprise, due largely to the difficulty of finding evidence bearing on one's hypothesis. (See Lewontin 1998 and Sterelny 2003:95-116.)

It's worth noting, however, that linguistic nativism is just one possibility for how Nature got language up and running. Just as it may be that a language faculty embodying knowledge of UG was somehow encoded in the human genome, it's also possible that that our ability to learn a language is based on a congeries of pre-existing competences, none of which is (or was initially — see below) specialized for language learning. Tomasello's theory of language acquisition, discussed above (§2.2.1.b), invites this alternative evolutionary perspective. On his view, the fundamental skills with which linguistic competence is acquired are skills that originally served, and still continue to serve, quite different, non-linguistic functions. For example, he argues that children's early word and phrase learning rest in part on their ability to share attention with others, to discern others' communicative intentions, and to imitate aspects of their behavior. There is reason to think that these abilities evolved independently of language, at least initially: imitation learning enabled the fast and high-fidelity transfer of learned skills between generations (see Tomasello 1999, 2000) and the ability to form beliefs about the mental states of others ('mind-reading' or 'theory of mind') enabled highly intelligent animals, such as our hominid

ancestors, to negotiate a complex social environment made up of similarly intelligent conspecifics. (See, e.g., Sterelny 2003.) On this sort of view, the ability to learn language piggy-backed on other capacities, which originally evolved for other reasons and which continue to serve other functions in addition to their linguistic ones.

You might wonder, however, whether this latter kind of account really differs substantively from that of a nativist. Assuming that she does not reject adaptationism altogether, the nativist will presumably be committed to the idea that the innate language organ, or faculty embodying knowledge of UG, was derived from pre-existing structures that were either functionless or had non-linguistic functions. These structures subsequently acquired linguistic functions through being selected for that reason: they became adaptations for language. But so too would the various capacities postulated by Tomasello. As soon as they started being used for language learning, that's to say, they would have been selected for that function (in addition to any other functions they might serve, and always assuming that linguistic abilities were on balance beneficial). Hence they too will over time become adaptations for language. On both Tomasello's and the nativist's view, in other words, the inborn structures responsible for language acquisition will have acquired the biological function of enabling language acquisition: they will be specialized for that purpose. Is Tomasello, then, a nativist?

No. First, even though the psychological abilities and mechanisms that Tomasello posits have been selected for linguistic functions, these abilities and mechanisms have continued to be used (and, plausibly, selected) for non-linguistic purposes, such as face recognition, theory of mind, non-linguistic perception, etc. So, whereas a central tenet of linguistic nativism is its insistence that the structures responsible for language learning are *task-specific*, Tomasello sees those structures as being much more general-purpose. In addition, and this is a second reason not to count Tomasello as a nativist, the inborn structures he posits are not plausibly interpreted as containing any kind of language specific information or representations. Yet a commitment to the role of inborn, language-specific information (such as knowledge of UG) is another hallmark of linguistic nativism.

Several theorists (e.g., Clark 1996, Tomasello 1999, and Sterelny, 2003) have stressed that in addition to working on human linguistic abilities directly, via changes to the parts of the genome coding for those abilities, natural selection can also bring about such changes indirectly, by making sure that our minds are embedded in certain kinds of environments. All sorts of animals create environments for themselves: this is called 'niche construction.' (The term is due to Odling-Smee, Laland and Feldman 1996.) Many animals also (or thereby) create environments for their offspring as well. And as Odling-Smee *et al*. 1996, Avital and Jablonka 2000, and Sterelny 2003 stress, animals' dispositions to modify the environments of both themselves and their offspring in certain ways are just as much potential objects of selection as are other of their traits.

To see this, suppose that an organism O has a genetically encoded disposition N to build a special kind of nest; suppose further that being raised in this kind of nest causes O-type offspring to have characteristic C; and suppose finally, that Os with C enjoy greater reproductive success than those without. Then, assuming that there is variation in N in the population, natural selection can operate so as to increase the proportion of Os with N — and hence also those with characteristic C — in the population. Down the track, Os will have C not by virtue of acquiring a special, genetically-encoded disposition-for-C. Rather, they will have C because their parents have the genetically-encoded disposition N, and Os whose parents have N 'automatically' develop C.

This toy example illustrates a further route by which language might have evolved in human beings. In addition to creating inborn language-learning mechanisms in individuals, natural selection may also have created dispositions to construct particular kinds of linguistic learning environments *in their parents*. For example, as Clark (1996) and Sterelny (2003) both speculate, Mother Nature might have worked on our dispositions to use 'Motherese' to our children, and/or on our tendency to talk about things that are current objects of the child's perceptual attention, in order to create learning environments conducive to the acquisition of language.

In principle, the existence of this sort of 'niche construction' can be accepted by all parties to the nativism controversy. That is, both Tomasello and Chomsky could agree that dispositions to construct 'linguistic niches' — environments in which languages are easy for human offspring to learn — may have been selected for in our species. Nevertheless, the notion of niche construction militates against the nativist, particularly when one takes into account the related notion of 'cumulative downstream niche construction.'

Cases of what Sterelny (2003: 149ff) calls 'cumulative downstream niche construction' occur when a generation of animals modifies an environment that has already been modified by earlier generations. A mountain thornbill's nest is an instance of downstream niche construction (since its offspring are affected by the thornbill's efforts). However, the construction is not cumulative, since the nest is built anew each year. By contrast, a rabbit warren extended and elaborated over several generations is an instance of cumulative construction: successive generations of offspring inherit an ever-more-complex niche and their other behaviors are tuned accordingly in ever-more-complex ways. Tomasello, 1999 and Sterelny 2003 stress that niche construction, including downstream niche construction, is not limited to the physical world: animals make changes to their social and epistemic worlds as well. For instance, chimpanzees live in groups (= construction of a social niche) and dogs mark their territory (= a change in their epistemic niche, relieving them of the necessity of remembering where the boundaries of their territory are). Humans, says Sterelny, echoing a theme of Tomasello 1999, are niche constructors "with a vengeance" (2003:149) and many of the changes they make to their physical, social and epistemic environments accumulate over many generations (think of a city, a democracy, modern science, a natural language). Such cumulative modifications allow for what Tomasello calls a "ratchet effect": a "cycle in which an improvement is made, becomes standard for the group, and then becomes a basis for further innovation." (Sterelny 2003: 150-1)

The idea of cumulative niche construction has obvious application to the case of language. If parents shape the linguistic environment of their offspring, and if we all shape the linguistic environments of our conspecifics (merely by talking to them!) then the possibility of a 'linguistic ratchet effect' is clearly open. Small changes made to the language of the group by one generation — changes which perhaps make it easier to learn, or easier to understand or produce — will be transmitted to later generations, who may in turn make further changes geared to increasing language learnability and ease of use. This scenario raises the possibility, already mentioned at the end of the last section, that language may have evolved so as to be learnable and usable by us, in addition to the converse scenario (stressed in much work on the evolution of language) that *we* had to change in many and complex ways in order to learn and use a language. Thus, we might speculate, languages' phonetic systems evolved so as to be congenial to our animal ears; their expressive resources (in particular, their vocabularies) evolved so as to fit our communicative needs; and perhaps, as Clark 1997 has suggested and as Tomasello 2003 implicitly takes for granted, natural language syntax evolved so as to suit our pre-existing cognitive and processing capacities. To be sure, the languages we have coded in our heads look complex and weird to linguists and psychologists and philosophers who are trying to put together theories about them. But, if languages and human minds have evolved in tandem, as surely they have, then languages may not look weird at all from the point of view of the brains that implement and use them.

All of these processes have likely played a role in the evolution of our capacities to learn and use a natural language. Pre-existing psychological, perceptual and motor capacities would have been recruited for the task of language learning and use. These capacities would have been honed and specialized further by natural selection for the performance of linguistic tasks. The functions of some of them, perhaps, would have become so specialized for language-related tasks that they cease to perform any non-linguistic functions at all — and to this extent, perhaps, linguistic nativism would be vindicated. At the same time, however, language itself would have been evolving so as the better to suit our cognitive and perceptual capacities, and our communicative needs. Given the fact that many different perceptual, motor and cognitive systems are implicated in language use and learning, and given the co-evolution of our minds and our languages, the truth about language evolution, when it emerges, is unlikely to be a simple. For this reason, it is unlikely to vindicate the nativist's notion that a specialized and monolithic 'language organ' or 'faculty' is at the root of our linguistic capacities.

Before leaving the question of language evolution, it is necessary to mention a recent paper by Hauser, Chomsky and Fitch 2002 on this topic. First, they distinguish (2002:1571) what they call the 'faculty of language in the narrow sense,' or 'FLN,' from the 'faculty of language in the broad sense,' or 'FLB.' The FLN is the "abstract linguistic computational system alone…which generates internal representations and maps them into the sensory-motor interface by the phonological system, and into the conceptual-intentional interface by the (formal) semantic system." (*Ibid*.) The FLB includes the FLN plus all the other systems (motor systems, conceptual systems, perceptual systems, and learning skills) which contribute to language acquisition and use.

Next, Hauser *et al*. speculate that the only thing that's really special about the human FLB is the FLN. That is, with the exception only of the FLN, FLB comprises systems that are shared with (or only slight modifications of) systems in other animals. Consequently, there is no mystery (or no more mystery than usual) about how these language-related abilities evolved. FLN, on the other hand, is distinctive to humans and what is special about it is its power of *recursion*, that is, its ability to categorize linguistic objects into hierarchically organized classes, and (on the behavioral side) for the generation of infinitely many sentences out of finitely many words. According to Hauser *et al*., the only real evolutionary mystery about language is how this capacity for recursion evolved — and this question, argue Hauser *et al*, is eminently addressable by normal biological methods (e.g., comparative studies to determine possible precursor mechanisms, etc.).

However, there are two difficulties with this scenario. First, there is evidence that the power of recursion posited by Hauser *et al*. as being distinctive of the human FLN is in fact not distinctive to humans, because it is not species specific. (See Esser, *et al*. 1997 and McGonigle, Chalmers and Dickinson, 2003.) Second, recursiveness is not language specific either, but is a feature of other domains of human cognition and endeavor as well. Our conceptual space, for instance, appears to be hierarchically ordered (poodles are a kind of dog, which are a kind of quadruped, which are a kind of animal, etc.). Similarly, the planning and execution of non-linguistic actions seems often to involve the sequencing and combining of smaller behavioral units into larger wholes. Recursion might well be an important part of the human language faculty, but it's apparently not specific either to us or to that faculty. Or, to put the point more bluntly: if it's Chomsky's view that recursiveness is the pivotal feature of the language faculty, and if recursiveness is a feature of human cognition and action more generally, then it's not clear that Chomsky remains a linguistic nativist.[26]

## 3.5 Pidgins and Creoles

It has been argued (by, e.g., Bickerton 1981, and Pinker, 1994:32-9) that the process by which a pidgin turns into a creole provides direct evidence of the operation of an innate language faculty. Pidgins are rudimentary communication systems that are developed when people speaking different languages come together (often in a commercial setting or when one people has conquered and is exploiting another) and need to communicate about practical matters. Creoles arise when pidgins are elaborated both syntactically and semantically, and take on the characteristics of *bona fide* natural languages.

Bickerton and, following him, Pinker, argue that creolization occurs when children take a pidgin as the input to their first language learning, and urge that the added complexity of the creole reflects the operation of the child's inborn language faculty. Moreover, they argue, since creole languages all tend to be elaborated in the same ways, and since they all respect the constraints of UG, the phenomenon of creolization also supports the idea that the inborn contribution to language acquisition is not just some general drive for an effective system of communication, but rather knowledge of linguistic universals.

There are two problems with this 'language bioprogram hypothesis,' as it is known in the creolization literature. The first concerns the claim (e.g., Bickerton 1981:43-70) that even creoles that developed in quite different areas of the world, and in complete isolation from one another, bear "uncanny resemblances" (Pinker 1994:25) to each other, not just in respecting the constraints of UG, but — even more surprisingly — in using fundamentally the same means to elaborate their root pidgins (e.g., in using the same syntactic devices to mark tense, aspect and modality). The stronger claim made by Bickerton — that Creoles use the same devices for the same grammatical purposes — is simply not true. For example, as Myhill (1991) argues, Jamaican Creole, Louisiana Creole, Mauritian Creole and Guyanese Creole mark tense, aspect and modality in ways that are quite different from those that Bickerton (1981) proposed as universal. (See, however, Mufwene 1999 for a case that confirms Bickerton's predictions.) The weaker claim — that creoles respect the constraints imposed by UG — has not, so far as I know, been contested. So we will assume, in what follows, that creoles, like other NLs, respect UG. The important question for our purposes is: how does this come about?

The bioprogram hypothesis claims that creolization occurs as a result of the action of the language faculty: children who learn language from degraded (e.g., pidgin) inputs are compelled by their innate knowledge of grammar to produce a fully-fledged natural language (the creole) as output. As an example of how children add

UG-constrained structure to languages learned from degraded inputs, Pinker cites the case of Simon, a deaf child studied by Newport and her colleagues, who learned American Sign Language (ASL) from parents who themselves were not exposed to ASL until their late teens. Although they used ASL as their primary language, Simon's parents were "in many ways…like pidgin speakers," says Pinker (1994:38). For instance, they used inflectional markers in an inconsistent way and often failed to respect the structure-dependence of the rules governing topicalization in that language.[27] But "astoundingly," says Pinker, "though Simon saw no ASL but his parents' defective version, his own signing was far better ASL then theirs…Simon must somehow have shut out his parents' ungrammatical 'noise.' He must have latched on to the inflections that his parents used inconsistently, and interpreted them as mandatory." (1994:39) Pinker views this as a case of "creolization by a single living child" (*ibid*.) and explains Simon's conformity to ASL grammar in terms of the operation of his innate language faculty during the acquisition period.

In a recent overview of the Simon data from the last 10 or so years, however, Newport 2001, stresses a number of facts that Pinker's presentation obscures or downplays. First, Simon's performance was not that of a native signer, although he did develop "his own version of ASL whose structure was more like that of other natural languages [than that of his parents' ASL]" (Newport 2001:168). For instance, Simon's morphology stabilized at a level that was "not as complex as native ASL" and he didn't acquire standard classifier morphemes if they were not used by his parents (*ibid*.). Secondly, Simon's success in learning a given rule seemed to vary with how well or badly his parents signed. For instance, Simon's parents used the correct inflectional morphology 60-75% of the time for a large class of verbs of motion, and in this case, Simon's own use of such morphology was 90% correct. However, some members of the class of classifier morphemes were correctly used by the parents only 40% of the time, and in this case, although Simon's performance was better than his parents', it was not at native signer level.

Newport argues that Simon appears to be 'cleaning up' his parents' language, that is, "bas[ing] his learning heavily on his input, but reorganiz[ing] this input to form a cleaner, more rule-governed system than the one to which he was exposed." (2001:168) She agrees that this result could be due to constraints imposed by an innate language faculty, but argues that it is also consistent with the existence of some more generalized propensity in children to generate systematic rules from noisy inputs, rightly pointing out that the latter hypothesis cannot be ruled out in advance of empirical test. (In this context, she notes (p.170) some preliminary studies suggesting that inferring systematic rules from messy data may indeed be a more general feature of learning in young children (though, interestingly, not in adults), for they can be seen to exhibit this tendency in non-linguistic pattern-learning contexts too.) Newport concludes that "the contrasts between Simon and his parents are in certain ways less extreme, and more reorganizational, than might be suggested by the Language Bioprogram Hypothesis…[H]e does not appear to be creating an entirely new language from his own innate specifications; rather, he appears to be following the predominant tendencies of his input, but he sharpens them, extends them, and forces them to be internally consistent." (2001:173).

If Newport *et al*. are right, the case of Simon does not seem to give much support to the nativist hypothesis. Moreover, the argument from creolization suffers a number of additional flaws. First, the Bickerton-Pinker view, which assigns a dominant role to child language learners in the creation of creoles, is but one of three competing hypotheses currently being explored in the creolization literature. According to the 'superstratist' hypothesis, creolization occurs not when children acquire language from pidgins, but when successive waves of adult speakers try to learn the language of the dominant culture as a second language. (Chaudenson 1992, for instance, defends this view about the origins of French creoles.) On this view, the additional devices seen in creoles are corruptions of devices seen in the dominant language. According to the 'substratist' hypothesis, creoles are again created by second language learners, rather than children, only the source of added structure is the first language of the learner. (Lumsden 1999 argues that numerous traces of a variety of African languages in Haitian creole support this hypothesis.) One need not take a stand on which of these views is correct in order to see that these competing explanations of creolization undermine Bickerton and Pinker's 'bioprogram' hypothesis. If creoles arise out of the attempts of adult learners to learn (and subsequently pass on to their children) another, non-native language, then what one might call 'contamination of the stimulus,' rather than the influence of an inborn UG in the learner, is what accounts for the UG-respecting ways in which creoles are elaborated.

However, there is a case of creolization in which these other hypotheses apparently fail to gain purchase, as Pinker (1994:37ff.) emphasizes. This is the case of the development of Idioma de Signos Nicaragüense (ISN, Nicaraguan Sign Language), a brand-new natural sign language which first emerged around 30 years ago in schools for the deaf in and around Managua. These schools were first set up in the 1970s, and ISN evolved from the hodge-podge of homesign systems used by students who entered the schools at that time. ISN is an interesting test case of the bioprogram hypothesis for two reasons. First, homesign systems are idiosyncratic and possess little syntactic structure: the natural-languagelike syntax of ISN could therefore not derive from substrate influence. And Spanish, the only potential candidate for superstrate influence was allegedly inaccessible to signers because of its auditory modality. Pinker claims that ISN provides another example of creolization and the workings of the innate language faculty: it is "created…in one leap when the younger children were exposed to the pidgin singing of the older children." (1994:36-7)

In their discussion of the development of ISN, however, Kegl, Senghas and Coppola (1999) show that things are not quite this straightforward. ISN did not develop 'in one leap,' from the very rudimentary homesigns or 'Mimicas' spoken by individual deaf students. Instead, its evolution was more gradual and was preceded by the creation of what Kegl *et al*. call "Lenguage de Signos Nicaragüense" (LSN), a kind of "pidgin or jargon" (181) that "developed from the point when these homesigners came together in the schools and began to share their homesigns with each other, quickly leading to more and more shared signs and grammatical devices" (180). In addition, the signers had access to Spanish language dictionaries, and their language was also influenced by the signing of Spanish-speaking, non-deaf teachers at the schools — signing which likely incorporated such grammatical devices of the teachers' language as were transferable to a non-vocal medium. (K. Stromswold, private communication.)

While Kegl *et al*. endorse the language bioprogram hypothesis that ISN emerged 'in one leap,' in the minds of children exposed to degraded Mimica or LSN inputs, their data are equally consistent with the idea that ISN developed more gradually by means of successive elaborations and innovations among a community of highly-motivated (because language-starved) young users. Indeed, as Kegl *et al*. themselves describe the history (p.187), this is precisely what happened. First, a group of signers, each with his or her own idiosyncratic form of Mimicas, entered the schools. Members of this group gained in expressive power as their individual Mimicas were enriched by borrowings from others' homesign systems. Then, a new cohort of Mimicas signers entered the schools. Their sign systems benefitted both from exposure to the Mimicas of their peers and from exposure to the richer system developed by the earlier group. Through a process of successive elaborations in this manner, LSN developed and then, by a similar series of steps, ISN developed. At present, all three sign systems are still being used in Nicaragua, presumably reflecting the different ages at which people are exposed to language and the kinds of inputs (ISN or LSN vs. signed and written Spanish or lipreading in regular schools) they receive. In addition, ISN and to a lesser extent LSN are still constantly changing — as one would expect if ISN were a community-wide work in progress, not the finished product of an individual child's mind,

## 3.6 Developmental Language Disorders and the Search for 'Language Genes'

Dissociations of language disorders acquired in adulthood (e.g., Broca's and Wernicke's aphasia) may tell us something about how language is organized in the mature brain, but cannot tell us much about how language is acquired or the role of innate knowledge in that process — a fact that nativists about language generally acknowledge. By contrast, language dissociations arising during childhood are sometimes held to bear strongly on the question of whether language is innate. Pinker (1994:297-314) articulates this latter line of thought, arguing that there is a double dissociation between 'general intelligence' and language in two developmental disorders called Williams Syndrome (WS) and Specific Language Impairment (SLI). People with WS have IQs well below the normal range (50-60), yet are able to speak fluently and engagingly about many topics. Those with SLI, by contrast, have normal ($\approx$90) non-verbal intelligence but speak effortfully and slowly, frequently making errors in their production and comprehension of sentences and words. Pinker argues that there is a double dissociation here, and that it supports the view that there is a special 'language acquisition device' that is separable from any general learning abilities children might possess. In addition, following Gopnik 1990a,b, and Gopnik and Crago 1991, he urges that the fact that the dissociation appears to concern aspects of syntax in particular indicates that the language faculty in question is the grammar faculty. Finally, and again following

Gopnik, he argues that since SLI appears to run in families and, in at least one case, displays a Mendelian inheritance pattern, what we have here is evidence not just of a 'grammar faculty,' but of a 'grammar gene.'

### 3.6.1 Williams Syndrome

WS is a rare genetic disorder with a complex phenotype. Physically, WS individuals display dismorphic facial features, abnormal growth patterns, gastrointestinal problems, early puberty, neurological abnormalities (including hypotonia, hyperreflexia, hyperacuisis and cerebellar dysfunction), defective vision and eye development, bad teeth, connective tissue abnormalities, and heart problems. Psychologically, in addition to their low non-verbal IQ and comparatively spared language abilities, they display relatively good audiovisual memory but very impaired visual-spatial abilities, leading to difficulties in daily life (e.g., getting dressed). They have outgoing personalities and are highly sociable to the point of overfriendliness, but also display numerous behavioral and emotional problems (especially hyperactivity and difficulty concentrating in childhood, and anxiety in later life). (Morris and Mervis 2000; Mervis *et al*. 2000.)

As to their language, there is currently a debate in the literature with regard to its normalcy. According to one point of view, the linguistic competence of WS individuals is remarkably normal, especially in comparison with that of similarly retarded individuals, such as those with Downs syndrome (Pinker 1994, 1997; Clahsen and Almazan 1998; Bello *et al*. 2004; Bellugi *et al*. 1998). While rather unusual in their choice of words (e.g., producing *chihuahua, ibis* and *condor* in addition to more usual animal words in a word fluency test) and despite an excessive use of clichés and social stock-phrases, their ability to use language, especially in conversational contexts, is more or less intact. For example, they may appear relatively normal in social interactions, and their processing of conditional questions and ability to repeat sentences with complex syntax is closer to that of normal controls than to matched Downs syndrome controls (Bellugi *et al*. 2000: 13, 15).

According to another school of thought, however, the language abilities of WS subjects might be more normal than those of Downs syndrome individuals, and might look remarkable in contrast to their own marked disabilities in other areas, but nonetheless display a number of abnormal characteristics across a variety of measures when investigated further. WS language shows "massively delayed" early acquisition, especially of vocabulary (Bellugi *et al*. 2000:11) and grammatical morphemes (Caprirci *et al*. 1996); overregularization of regular plural and past tense endings as well defective competence with regard to irregular nouns and verbs (Clahsen and Almazan 2001); "inordinate difficulty with morphosyntax" (Morris and Mervis 2000: 467; see also Volterra *et al*. 1996; Karmiloff-Smith *et al*. 1997; Levy and Hermon 2003); and impaired mastery of relative clause constructions (Grant *et al*., 2002), embedded sentences, and (in French) grammatical gender assignment (Karmiloff-Smith *et al*. 1997). Indeed, Bellugi *et al*., 2000, found that WS children's performance on a sentence-repetition task was indistinguishable from that of matched controls diagnosed with Specific Language Impariment, or SLI (see below, §3.6.2). Findings such as these lead experts such as Annette Karmilloff-Smith to urge "dethroning the myth" of WS' "intact" syntactic abilities (Karmiloff-Smith *et al*. 2003) and move Ursula Bellugi — formerly a proponent of the 'spared language' viewpoint — to caution that "because their language abilities are often at a level that is higher than their overall cognitive abilities, individuals with WMS might be perceived to be more capable than they really are." (Bellugi *et al*. 1999.)

In contrast to its cognitive profile, which is, as we have seen, a subject of debate, the genetic basis of WS is known. It results from a ≈1.5 Mb deletion encompassing the elastin gene ELN at chromosome 7q11.23; most cases appear to be due to new mutations. ELN is crucial in synthesizing elastin, a protein which holds cells together in the elastic fibers found in connective tissues throughout the body and in especially high concentrations in cartilege, ligaments and arterial walls. Failure to synthesize this protein disrupts development in numerous ways, from the first trimester onwards, and gives rise through processes that are not well understood to the raft of symptoms associated with the syndrome. (Morris and Mervis 2000; Mervis *et al*. 2000.)

### 3.6.2 Specific Language Impairment

In contrast with Williams syndrome, in which one sees comparatively spared language in the face of mild to moderate mental retardation and numerous physical defects, specific language impairment ('SLI') is diagnosed

when (i) non-verbal intelligence as measured by standard IQ tests is normal; (ii) verbal IQ is well below normal; and (iii) obvious causes of language impairment (e.g., deafness, frank neurological damage) can be ruled out. As one might expect given these diagnostic criteria, a diagnosis of SLI embraces a highly heterogeneous collection of language-related deficits, not all of which co-occur in every case of language impairment. (Bishop, 1994; Bishop *et al*,. 2000.) These include:

- productive and receptive phonological deficits (e.g., difficulty producing clusters of consonants, as in *spectacle,* and failure to show categorical perception of phonemes differentiated by place of articulation (/ba/ vs. /ga/) and voicing (/ba/ vs. /pa/);
- morphological deficits (e.g., generation of past tenses or plurals by using affixes);
- productive and receptive syntactic deficits (e.g., analyzing 'reversible' passives (*Katie kissed Jacob* vs. *Jacob was kissed by Katie*), complex dative constructions (e.g., *Katie gave Jacob the book*) and anaphora (e.g., *Katie said that Sarah scratched her* vs. *Katie said that Sarah scratched herself).*

As a consequence of this heterogeneity, SLI, researchers have introduced a number of subtypes of the disorder, including such things as 'Verbal auditory agnosia,' 'Lexical-syntactic deficit syndrome' and 'Phonological programming syndrome' (Bishop 1994). Also as a consequence, and in part because studies do not always distinguish between different subtypes, the etiology of SLI in general is not well understood (O'Brien *et al*. 2003), although recent research suggests at least two distinct genetic loci are involved in at least some subtypes of the disorder (Bishop 2006). Some posit an underlying defect in the 'grammar module.' For instance, Rice and Wexler (1996) attribute SLI individuals' morphological deficits to a missing UG principle, namely, the principle of inflection, and Van der Laly and Stollwerk 1997, attribute some SLI children's difficulty with anaphora to their failure to acquire Binding Theory. Others see non-linguistic defects, such as auditory, memory or processing deficits as the root problem. For instance, Tallal 1980, 1985 argue that many SLI cases result from deficits in the processing of rapid auditory stimuli, giving rise to a failure to learn to distinguish phonemes correctly, which in turn leads to a failure to acquire other aspects of grammar. Others, such as and Norbury, Bishop and Briscoe 2002 argue that such children's limited processing capacities are the culprit.

While the varied symptomatology of SLI suggests that no unified theory of its etiology might be forthcoming, the cause of the disorder is comparatively well understood in the case of one subtype, involving a severe disruption of morphosyntax (i.e., the rules governing the formation of words from smaller semantic units, or morphemes). This subtype, seen in about half the members of a large, three-generation English family, the KE's, and in another, unrelated individual, has been traced to a specific genetic mutation, the function of which is actively under investigation.

The KE family has received much press since the early 1990s, when Gopnik 1990a,b and Gopnik and Crago 1991 (see also Gopnik 1997) proposed that their morphosyntactic deficits were caused by a mutation in a single dominant gene normally responsible for the encoding of grammatical features, such as function words and the inflections used to mark number, tense, aspect, etc. According to Gopnik, the affected KE's are 'feature blind' as a consequence of this mutation. And according to Pinker (1994), their pedigree (Fig. 6) and specifically morphosyntactic deficits constitutes "suggestive evidence for grammar genes … genes whose effects seem most specific to the development of the circuits underlying parts of grammar" (Pinker 1994:325).
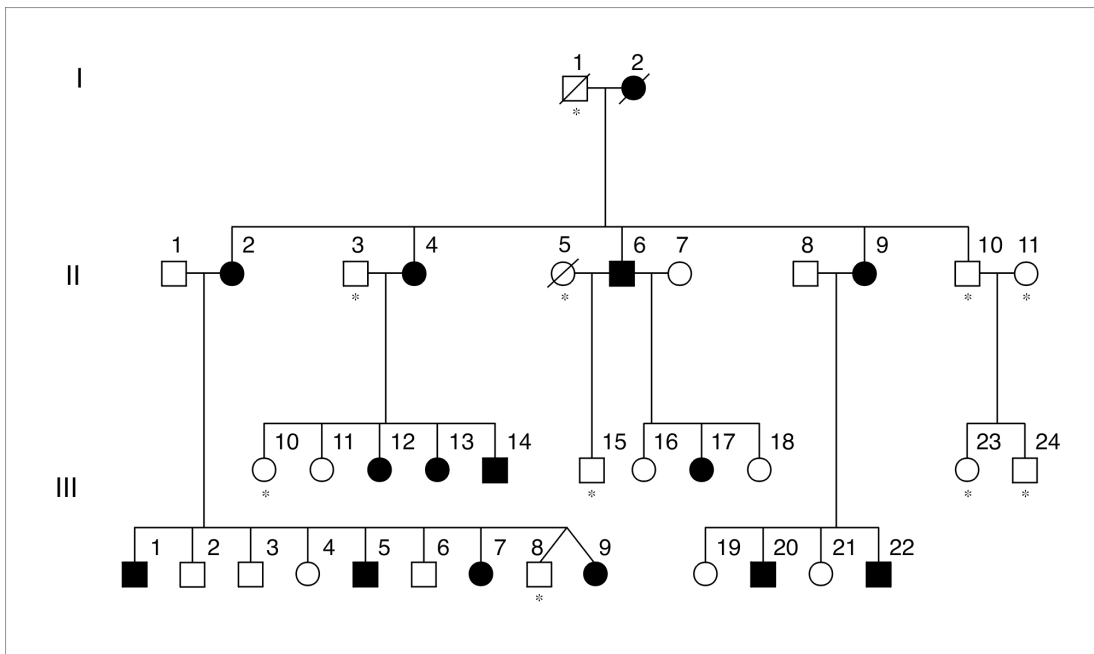
Figure 6. The KE family pedigree
(*Image used by permission of Simon E. Fisher*)

Other intensive studies of the KE family, by Vargha-Khadem and colleagues (e.g., Vargha-Khadem *et al* 1995, 1998; Watkins *et al*., 2002) have vigorously disputed the hypothesis that the root cause of the KE's language disorder is a syntactic deficit. Instead, they argue, the KE phenotype is much broader than Gopnik's account suggests, and their 'feature blindness' is merely one among the many effects of an underlying articulatory problem. As characterized by Vargha-Khadem's team, the affected KE's speech is effortful, "sometimes agrammatical and often unintelligible" (Watkins *et al*. 2002:453), and shows impairments not just in morphosyntax (e.g., regular plural and past tense endings) but also in the formation of irregular past tenses (where correct usage is lexically determined, rather than rule-governed) and in sentence-level syntax, particularly word order. Comprehension, too, is impaired at the level of syntax as well as words, and as is their reading of both words and non-words. These results indicate that the KE's problems go beyond morphosyntax, and the fact that affected KE's have significantly lower non-verbal IQs (by 18-19 points; Vargha-Khadem *et al*. 1995) than unaffected family members indicates that their deficits may be further reaching still.[28] Finally, affected KE's have trouble sequencing and executing non-language-related face, mouth and tongue movements and show abnormal activation not just of speech but also of motor areas on fMRI scans (Liegeois, *et al*. 2003); this deficiency in 'orofacial praxis' supports Vargha-Khadem's hypothesis that the root problem for the KE's is articulatory.

As Gopnik noted, the pattern of inheritance in the KE family suggests that a single, dominant gene is responsible for the disorder. (See fig. 6) In the early 1990's, Fisher and colleagues began working to isolate the gene. First, it was localized to a region on chromosome 7q31 containing about 100 genes (Fisher *et al*. 1997, 1998; O'Brien *et al*. 2003). Later, it was identified (Lai *et al*. 2001; Fisher et al. 2003) as the gene FOXP2, which encodes a regulatory protein or 'transcription factor' (i.e., a protein that helps to regulate the rate of transcription of other genes in the genome — in the case of FOXP2, the protein acts to inhibit transcription of the downstream gene(s)). In affected family members, a single base-pair substitution in the gene coding for this regulatory protein leads to the insertion of the amino acid arginine (instead of the normal histamine) in an area of the protein (viz., the 'forkhead binding domain') that is critical for its ability to modulate the transcription of the downstream DNA. As a consequence, FOXP2 cannot perform its normal regulatory role in affected KE family members.

The failure of FOXP2 to perform its normal role in turn leads to abnormal brain development in affected KE individuals. Studies of other animals and humans (e.g., Lai *et al*. 2003; Takahashi *et al*., 2003; Ferland *et al*. 2003; Teramitsu *et al*. 2004) show that FOXP2 is normally highly expressed in both development and adulthood

in two distinct brain circuits, One is a corticostriatal circuit, in which inputs from the prefrontal and premotor cortex are modulated by the basal ganglia and the thalamus, and then sent back to prefrontal and premotor cotical areas; the other is an olivocerebellar circuit, in which sensory input is sent via the spinal cord for processing in the medulla, cerebellum and thalamus before being handed on to prefrontal cortex. (See fig. 7.) The basal ganglia are known to be involved in the sequencing and reward-based learning of motor behaviors (Graybiel 1995, 1998), and the cerebellar circuit, while less well understood, is thought to be a proprioceptive circuit involved in motor regulation and coordination (Lieberman 2002). FOXP2 is expressed in homologues of these areas in other species (e.g., canaries, zebra finches, rats) and in all species studied, these areas are involved in motor sequencing and coordination (Sharff and White 2004).
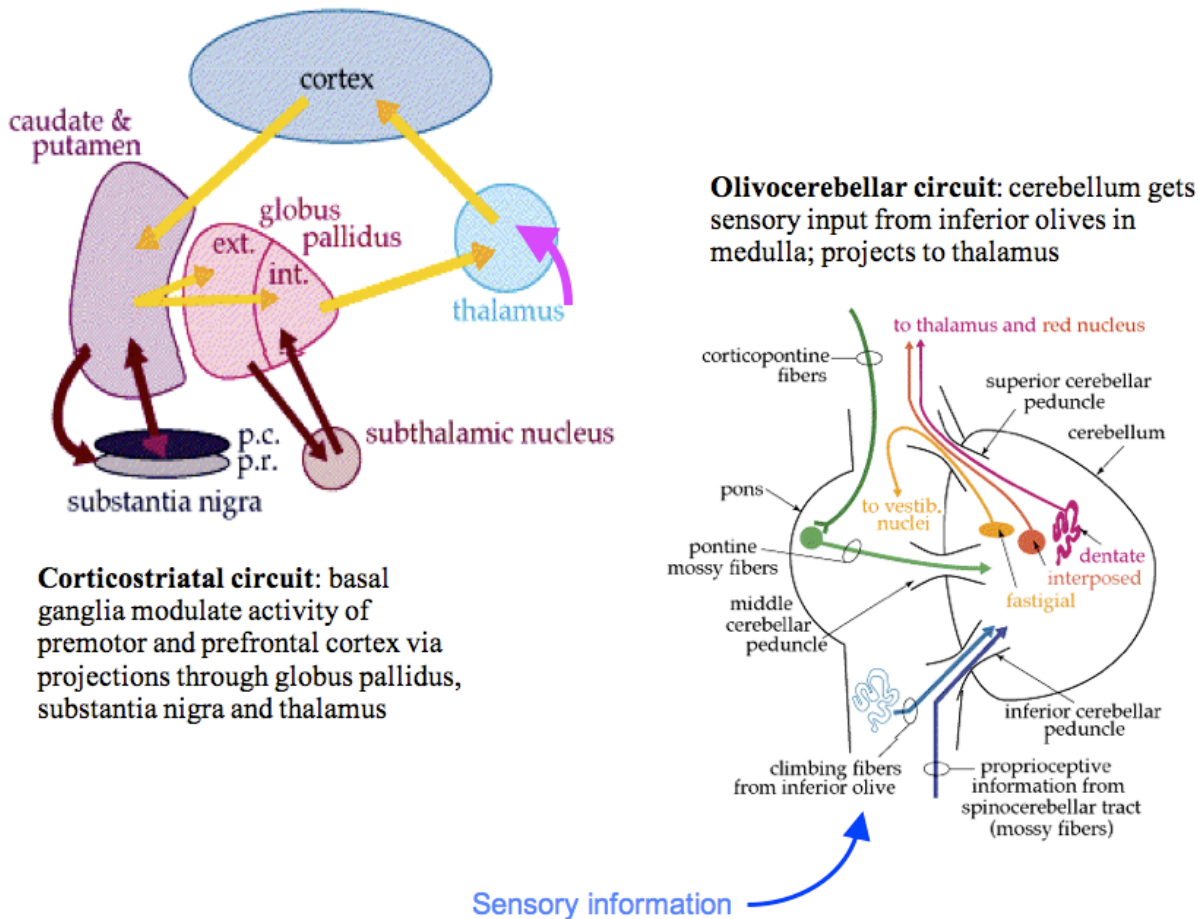


Figure 7. Two circuits in which FOXP2 is expressed. (Based on figures by Diana Weedman Molavi, The Washington University School of Medicine Neuroscience Tutorial).

So, what appears to be the case is that affected KE family members' language difficulties result from a mutation in the FOXP2 gene, which results in abnormal development of the striatal, cerebellar and cortical areas necessary for the sequencing and coordination of speech-related movements of the mouth, tongue and possibly larynx; MRI scans of affected family members showing reduced gray matter density in these areas support this hypothesis, as do fMRI scans showing abnormal striatal and cortical activation during receptive and active language processing (Belton *et al*. 2003; Liegeois, et al. 2003.)

Vargha-Khadem speculates (cf. Watkins *et al*. 2002:463) that those of the KE's deficits that do not appear to be motor related (e.g., their comprehension and reading problems, their difficulties with word order and syntax) are a result of impaired learning that itself results from their motor deficits. For instance, impaired articulation could lead to impoverished phonological representations, which would then impair the acquisition of morphological and morphosyntactic knowledge, which would then constitute a poor basis for further syntactic learning. Impaired representation at all of these levels would then express itself in receptive language and reading, as well as in the realm of spoken language. Another possible explanation of the KE's non-articulatory deficits, which is not necessarily in competition with the previous one, derives from the fact that the basal ganglia are also known

to be implicated in working memory (Bosnan 2004) and reward-based learning (e.g., classical conditioning) that is mediated by dopaminergic circuits that interact with basilar structures (Lieberman 2002). If reward-based learning and working memory are impaired in the KE's, then this could explain not only their higher-level syntactic deficits, but also their overall lower IQ (Lieberman 2002).

Neither of these explanations of the KE's seems especially congenial to the linguistic nativist. For both tacitly assume that language learning, including syntactic learning, is not (or not entirely) subserved by special-purpose mechanisms. Rather, it is mediated by more general motor circuitry (according to the Vargha-Khadem hypothesis) or reward-based learning and working memory abilities (according to Lieberman) that are also involved in other learning tasks.

On the other hand, however, there is evidence that FOXP2 is particularly implicated in vocal learning and expression. First, it is highly expressed in songbirds that modify their innate vocal repertoires: in canaries it is expressed seasonally, when adult birds modify their songs (Teramitsu *et al*. 2004) and in zebra finches, it is expressed more at the time when young birds learn their songs (Haessler *et al*. 2004). In addition, there is evidence that the variant of the FOXP2 gene that is present in humans has undergone strong positive selection in the hominid line (Enard *et al* 2002; Zhang *et al*. 2002). The protein produced by human FOXP2 differs in just three out of its 715 constituent amino acids from that of the mouse, and a recent analysis (Zhang *et al*. 2003) indicates that two of these differences are unique to the hominid lineage. According to Enard *et al* 2002, the fact that these two differences are fixed in the human genome, whereas no fixed substitutions occurred in the lineage of our closest relatives, the chimpanzees, suggests that those changes were strongly selected for in our lineage; Enard *et al*. put the date of fixation of these changes in the human population at around 200,000 years ago. This date accords well with at least some estimates of the emergence of modern human language, suggesting that the vocal capacities underwritten by FOXP2 — and impaired in those lacking the gene — are after all critical to language competence.

### 3.6.3 The grammar module and the genetics of language

At this point, two questions arise. First, is there a double dissociation between language capacities and general cognitive capacities to be found in a comparison of Williams syndrome and SLI? Second, what does our current knowledge of the role of FOXP2 in language development tell us about linguistic nativism?

As to the first question, there seems to be no double dissociation. First of all, WS individuals' language, while startling in contrast with their level of mental retardation, is not normal; indeed, as we have seen, it is indistinguishable on some tests from that of language impaired individuals. In addition, as Thomas and Karmiloff-Smith 2002, caution, it is not at all clear that one can assume, in the case of a pervasive developmental disorder like Williams syndrome, that apparently 'intact' competences are a result of normal development of the underlying neurological and psychological structures. That is, given the known capacity of the brain to compensate for deficits in one area by cobbling together a solution in another, one cannot assume that there is a 'language module' in WS patients which develops more or less normally despite other cognitive systems' being massively disrupted. Thomas and Karmiloff-Smith argue that the numerous discrepancies between WS language development and that of normal children suggests that this 'residual normality' assumption is misguided in this case, thus undermining the claim that what is spared in WS is 'the language (or grammar) module.'

Moving to the other side of the dissociation, since it is hard to say exactly what about language is disrupted in cases of SLI, it is difficult to determine whether this disruption is specific to language, let alone grammar. While researchers like Van der Lely and Christian 2000, and van der Lely and Ullman 2001 argue that there is a purely grammatical form of the deficit, which does support the hypothesis of a grammar module, this is controversial, as we have seen above. Certainly consideration of the KE's does not support such a hypothesis. Their root deficit appears to concern orofacial praxis, rather than language specifically; and in addition, their 'general intelligence,' as measured by tests of non-verbal intelligence, while "normal," nonetheless appears to have been affected by their neurological and/or linguistic abnormalities — witness their scores 18-19 points lower than those of their relatives. It is, in other words, unclear that there is any dissociation of language and general intelligence in this case at all. One can conclude that as things stand now, SLI seems to be so heterogeneous a disorder as to defy

neat characterization, and that consideration of this disorder does not support the view that there is a language or grammar module that functions independently of other cognitive processes.

The second question asked above was: what can be learned about the innate basis of language from a consideration of the KE's and FOXP2. In a recent article, Marcus and Fisher (2003) argue that the kinds of results discussed above offer valuable insights into the ways that language is implemented in the brain and controlled (to the extent that it is) by the genes. However, they refrain — rightly in my view — from drawing morals to the effect that FOXP2 is a "gene for language" or even "for articulation." The effects of FOXP2 are wider than this (it is expressed in the developing heart and lungs, in addition to the brain -- REF) and the functions of the neural circuits in which it is active are as yet too poorly understood to do more than gesture at the ways in which FOXP2 is involved in constructing the human linguistic phenotype.

All the topics covered in §3 deserve books of their own. My aim here has been to sketch the ways in which modern understanding of the mind reveals the inadequacy and implausibility of the claim that humans have innate representations of UG that are responsible for their acquisition of language. There are likely many, many processes implicated in the attainment of linguistic competence, that many of them are likely specialized by natural selection for linguistic tasks, but that many of them also retain their other, and older, functions. The linguistic nativist's theory views our acquisition of grammatical competence as a *simple* matter — one that can be described at one level of explanation, and in terms of a single kind of process. This is very unlikely to be the case. Multiple systems and multiple processes are at work in the acquisition of linguistic knowledge, and our understanding of language acquisition, when it comes, is likely to involve theories of many kinds and at many different levels, and to resemble the theory of the Chomskyan nativist in few or no respects.[29]

# Bibliography

- Adolphs, R. (1999). Social Cognition and the Human Brain. *Trends in Cognitive Science, 3*(12), 469-479.
- Avital, E. and Jablonka, E. (2000). *Animal Traditions: Behavioural Inheritance in Evolution*, Cambridge: Cambridge University Press,
- Ayoun, D. (2003). *Parameter Setting in Language Acquisition*, London: Continuum International Publishing Group, Incorporated.
- Baker, C.L. (1979). "Syntactic Theory and the Projection Problem", *Linguistic Inquiry*, 10: 533-81.
- Bates, E. and Roe, K. (2001). "Language development in children with unilateral brain injury", in *Handbook of Developmental Cognitive Neuroscience*, C.A. Nelson and M. Luciana,(eds.), Cambridge: MIT Press, pp. 281-307.
- Beale, J.M. and Keil, F.C. (1995). "Categorical Effects in the Perception of Faces", *Cognition*, 57: 217-239.
- Bello, A., Capirci, O., & Volterra, V. (2004). Lexical production in children with Williams syndrome: spontaneous use of gesture in a naming task. *Neuropsychologia, 42*(2), 201-213.
- Bellugi, U., & Lai, Z. (1998). Neuropathological and cognitive alterations in Williams syndrome and Down syndrome. *Faseb Journal, 12*(4), A355-A355.
- Bellugi, U., Lichtenberger, L., Mills, D., Galaburda, A., & Korenberg, J. R. (1999). Bridging cognition, the brain and molecular genetics: evidence from Williams syndrome. *Trends in Neurosciences, 22*(5), 197-207.
- Bellugi, U., L. Lichtenberger, *et al.* (2000). "The neurocognitive profile of Williams syndrome: A complex pattern of strengths and weaknesses." *Journal of Cognitive Neuroscience* 12: 7-29.
- Belton, E., C. H. Salmond, *et al.* (2003). "Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia." *Human Brain Mapping* 18(3): 194-200.
- Berwick, R.C. and Weinberg, A.S. (1983). *The Grammatical Basis of Linguistic Performance: Language Use and Acquisition*, Cambridge, MA: MIT Press.
- Berwick, R. C. (1998). Language evolution and the Minimalist Program: the origins of syntax. In J. R. Hurford, M. Studdert-Kennedy & C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 320-340). Cambridge: Cambridge University Press.
- Bickerton, D. (1981). *Roots of Language*: Karomer Publishers, Inc.

- ⸺ (1998). Catastrophic evolution: the case for a single step from protolanguage to full human language. In J. R. Hurford, M. Studdert-Kennedy & C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 341-358). Cambridge: Cambridge University Press.
- ⸺ (2000). *Language and Species*, Chicago : University of Chicago Press.
- Birdsong, D. and Molis, M. (2001). "On the evidence for maturational constraints in second-language acquisition", *Journal of Memory and Language*, 44: 235-49.
- Bishop, D. V. M. (1994). "Is Specific Language Impairment A Valid Diagnostic Category - Genetic And Psycholinguistic Evidence." *Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences* 346(1315): 105-111.
- ⸺, P. Bright, *et al*. (2000). "Grammatical SLI: A distinct subtype of developmental language impairment?" *Applied Psycholinguistics* 21(2): 159-181.
- Bohannon, J.N. and Stanowicz, L. (1988). "The issue of negative evidence: Adult responses to children's language errors", *Developmental Psychology*, 24: 684-89.
- Bohannon, J.N., MacWhinney, B. and Snow, C. (1990). "No negative evidence revisited: Beyond learnability or who has to prove what to whom", *Developmental Psychology,* 26: 221-26.
- Bohannon, J. N., Padgett, R. J., Nelson, K. E., & Mark, M. (1996). Useful evidence on negative evidence. *Developmental Psychology, 32*(3), 551-555.
- Bosman, C., R. Garcia, *et al*. (2004). "FOXP2 and the language working-memory system." *Trends In Cognitive Sciences* 8(6): 251-252.
- Botha, R. P. (2003). *Unravelling the Evolution of Language*. U.S.: Elsevier Science.
- Bradlow, A.R., Pisoni, D.B., Akahane-Yamada, R., and Tohkura, Y. (1997). "Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production", *Journal of the Acoustical Society of America*, 101: 2299-2310.
- Brandon, R.N. and Hornstein, N. (1986a). "From Icons to Symbols: Some Speculations on the Origins of Language", *Biology and Philosophy*, 1 :169-89.
- Brown, R. (1973). *A first language: The early stages*, Cambridge, MA: Harvard University Press.
- Brown, R., Cazden, C. and Bellugi, U. (1969). "The Child's Grammar from I to III", in *Minnesota Symposium on Child Psychology, Vol. II,* J. Hill (ed.), Minneapolis: University of Minnesota Press.
- Brown, R. and Hanlon, C. (1970). "Derivational Complexity and Order of Acquisition in Child Speech", in *Cognition and the Development of Language,* J. R. Hayes (ed.), New York: John Wiley and Sons.
- Buonomano, D.V. and Merzenich, M.M. (1998). "Cortical plasticity: From synapses to maps", *Annual Review of Neuroscience*, 21: 149-186.
- Campanella,. S., Chrysochoos, A., Bruyer, R. (2001). "Categorical perception of facial gender information: Behavioural evidence and the face-space metaphor", *Visual Cognition*, 8 (2): 237-262.
- Capirci, O., L. Sabbadini, *et al*. (1996). "Language development in Williams syndrome: A case study." *Cognitive Neuropsychology* 13(7): 1017-1039.
- Catania, A.C. (1990). "What good is 5 percent of a language competence", *Behavioral and Brain Sciences*, 13 : 729-730.
- Cavalli-Sforza, L.L. (1997). "Genes, peoples, and languages", PNAS: 94: 7719-772
- Chambers, K.E., Onishi, K.H., and Fisher, C. (2003). "Infants learn phonotactic regularities from brief auditory experience", *Cognition*, 87: 69-77.
- Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences, 10*(7), 335-344.
- Chaudenson, R. (1992). *Des îles, des hommes, des langues*. Paris: L'Harmattan.
- Chomsky, N. (1957). *Syntactic Structures*, The Hague : Mouton.
- ⸺ (1959). Review of Skinner's *Verbal Behavior*, *Language*, 35: 26-58.
- ⸺ (1965). *Aspects of the Theory of Syntax*, Cambridge, MA: MIT Press.
- ⸺ (1975). *Reflections on Language*, London : Fontana.
- ⸺ (1980). *Rules and Representations*. New York: Columbia University Press.
- ⸺ (1981). *Lectures on Government and Binding,* Hawthorne, NY: Walter De Gruyter Incorporated.
- ⸺ (1982). *The Generative Enterprise*, Dordrecht: Foris Publications.
- ⸺ (1986). *Knowledge of Language, Its Nature, Origin and Use*, NY: Praeger.
- ⸺ (1988). *Language and Problems of Knowledge, The Managua Lectures*, Cambridge, MA: MIT Press.
- ⸺ (1975a). *The Logical Structure of Linguistic Theory*, NY: Plenum.

- ——— (1980). *Rules and Representations*, NY: Columbia University Press.
- ——— (1990). "On the nature, acquisition and use of language", in *Mind and Cognition: A Reader,* W.G. Lycan (ed.), Cambridge MA and London UK: Blackwells, pp.627-45.
- ——— (1995). *The Minimalist Program,* Cambridge: MIT Press.
- Chouinard, M.M. and Clark, E.V. (2003) "Adult reformulations of child errors as negative evidence," *Journal of Child Language*, 30:637-69.
- Clahsen, H. and Muysken, P. (1986). "The availability of universal grammar to adult and child learners", *Second Language Research*, 2: 93-119.
- ——— and M. Almazan (1998). "Syntax and morphology in Williams syndrome." *Cognition* 68(3): 167-198.
- ——— (2001). Compounding and inflection in language impairment: Evidence from Williams Syndrome (and SLI). *Lingua, 111*(10), 729-757.
- Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*, Cambridge, MA: MIT Press.
- Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. Cambridge, Ma: Bradford Books/MIT Press.
- Cowen T. and Gavazzi, I. (1998). "Plasticity in adult and ageing sympathetic neurons", *Progress in Neurobiology*, 54:249-88.
- Cowie, F. (1997). The Logical Problem of Language Acquisition. *Synthese, 111*, 17-51.
- ——— (1999). *What's Within: Nativism Reconsidered*, New York: Oxford University Press.
- Crain, S. and Pietroski, P.M. (2001). "Nature, Nurture, and Universal Grammar", *Linguistics and Philosophy,* 24: 139-86.
- ——— (2002). "Why language acquisition is a snap", *Linguistic Review*, 19 (1-2): 163-183.
- Crain, S. (1991). "Language acquisition in the absence of experience", *The Behavioral and Brain Sciences*, 4: 597-650.
- Curtiss, S. (1977). *Genie: a Psycholinguistic Study of a Modern-day "Wild Child"*. New York: Academic Press.
- Dawkins, R. and Krebs, J.R. (1979). "Arms races between and within species", *Proceedings of the Royal Society of London, Series B — Biological Sciences*, 205: 489-511.
- DeCaspar, A.J. and Spence, M.J. (1986). "Prenatal maternal speech influences newborns' perception of speech sounds", *Infant Behavior and Development*, 9:133-150.
- DeGraff, D. (1999a). Creolization, Language Change, and Language Acquisition. In M. Degraff (Ed.), *Language Creation and Language Change: Creolization, diachrony, and development*. (pp. 1-46). Cambridge, MA: MIT Press.
- DeGraff, D. (Ed.). (1999b). *Language Creation and Language Change: Creolization, Diachrony and Development*. Cambridge, Ma.: MIT Press.
- Demetras, M.J., Post, K.N. and Snow, C.E. (1986). "Feedback to first language learners: The role of repetitions and clarification questions", *Journal of Child Language*, 13: 275-92.
- Descartes, R. (1984), "Discourse on the Method", in *The Philosophical Writings of Descartes*, Vol. II, J. Cottingham, R. Stoothoff and D. Murdoch (eds.), and trans. (1984), Cambridge: Cambridge University Press.
- Devitt M. (2006). *Ignorance of Language*, Oxford: Oxford University Press.
- Dick F., Bates E., Wulfeck B., Utman J.A., Dronkers N. and Gernsbacher M.A. (2001). "Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals", *Psychological Review*, 108: 759-788.
- Dronkers, N.F., Redfern, B.B. and Knight, R.T. (2000). "The Neural Architecture of Language Disorders", in *The New Cognitive Neurosciences, 2nd Ed.*, M.S. Gazzaniga (ed.), Cambridge: MIT Press, pp. 949-58.
- Dunbar, R. (1996). *Grooming, Gossip and the Evolution of Language*, Cambridge, Ma.: Harvard University Press.
- Eimas, P.D. (1975). "Speech perception in early infancy", in *Infant Perception, Vol.2. From sensation to cognition*, L. Cohen and P. Aalapatel (Eds.), New York: Academic Press. Pp.193-231.
- Elbert, T., Heim, S. and Rockstroh, B. (2001). In *Handbook of Developmental Cognitive Neuroscience,* C.A. Nelson and M. Luciana, Cambridge: MIT Press, pp. 191-202.
- Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D. and Plunkett, K. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*, Cambridge: Bradford Books/MIT

Press.

- Elman, J.L. (1998). "Generalization, simple recurrent networks, and the emergence of structure", in *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society,* M.A. Gernsbacher and S.J. Derry (eds.), Mahwah, NJ: Lawrence Erlbaum Associates.
- Enard, W., Przeworski, M., Fisher, S., Lai, C., Wiebe, V., Kitano, T., *et al*. (2002). "Molecular evolution of FOXP2, a gene involved in speech and language",. *Nature, 418*, 869-872.
- Epstein, S., Flynn, S., and Martohardjono, G. (1996). "Second language acquisition: theoretical and experimental issues in contemporary research," Brain and Behavioral Sciences, 19:677-758.
- —— (1998). "The strong continuity hypothesis: some evidence concerning functional categories in adult L2 acquisition". In The generative study of second language acquisition, S. Flynn, G. Martohardjono and W. O'Neil (eds.), Mahwah, NJ: Lawrence Erlbaum, pp.61-77.
- Esser, K-H., Condon, C.J., Suga, N., and Kanwal, J.S. (1997). "Syntax processing by auditory cortical neurons in the FM—FM area of the mustached bat Pteronotus parnellii", *Proc. Natl. Acad. Sci. USA*. 94: 14019—14024.
- Etcoff, N.L., Magee, J.J. (1992). "Categorical Perception of Facial Expressions", *Cognition* 44: 227-240.
- Farrar M.J. (1990). "Discourse and the acquisition of grammatical morphemes", *Journal of Child Language*, 17: 607-24.
- —— (1992). "Negative Evidence and Grammatical Morpheme Acquisition", *Developmental Psychology*, 28: 90-98.
- Feldman, J. (1972). "Some decidability results on grammatical inference and complexity", *Information and Control*, 20: 244-62.
- Ferland, R. J., T. J. Cherry, *et al*. (2003). "Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain." *Journal Of Comparative Neurology* 460(2): 266-279.
- Fisher, S. E., F. Vargha-Khadem, *et al*. (1997). "Localisation of a gene implicated in a severe speech and language disorder." *American Journal Of Human Genetics* 61(4): A28-A28.
- ——, F. Vargha-Khadem, *et al*. (1998). "Localisation of a gene implicated in a severe speech and language disorder," *Nature Genetics* 18(3): 298-298.
- ——, C. S. L. Lai, *et al*. (2003). "Deciphering the genetic basis of speech and language disorders." *Annual Review of Neuroscience* 26: 57-80.
- Flege, J.E., Yeni-Komshian, G.H. and Liu, S. (1999). "Age constraints on second language acquisition", *Journal of Memory and Language*, 41: 78-104.
- Fodor, J.A. (1981). "The Present Status of the Innateness Controversy", in *RePresentations: Philosophical Essays on the Foundations of Cognitive Science*, Cambridge: MIT Press/Bradford Books, pp. 257-316.
- Fodor, J.A. (1998). *Concepts: Where Cognitive Science went Wrong*, New York: Oxford University Press.
- Gold, E.M. (1967). "Language Identification in the Limit", *Information and Control*, 10: 447-74.
- Goldin-Meadow, S., and Mylander, C. (1983). "Gestural communication in deaf children: noneffect of parental input on language development", *Science*, 221: 372-4.
- Goldin-Meadow, S., and Mylander, C. (1990). "The role of parental input in the development of a morphological system", *Journal of Child Language*, 17(3): 527-63.
- Goldin-Meadow, S., and Mylander, C. (1998). "Spontaneous sign systems created by deaf children in two cultures", *Nature*, 391:279-81.
- Gopnik, M. (1997). "Language deficits and genetic factors", *Trends in Cognitive Science*, 1: 5-9
- —— (1990a) "Feature-blind grammar and dysphasia", *Nature*, 344:715.
- —— (1990b) "Genetic basis of a grammar defect", *Nature, 347: 26.
- ——, and Crago, M. (1991). "Familial aggregation of a developmental language disorder", *Cognition*, 39: 1-50.
- Grant, J., V. Valian, *et al*. (2002). "A study of relative clauses in Williams syndrome." *Journal Of Child Language* 29(2): 403-416.
- Graybiel, A. M. (1995). "Building action repertoires: Memory and learning functions of the basal ganglia." *Current Opinion In Neurobiology* 5(6): 733-741.
- Graybiel, A. M. (1998). "The basal ganglia and chunking of action repertoires." Neurobiology Of Learning And Memory 70(1-2): 119-136.
- Griffiths, P. (2002). "What is innateness?", *Monist*, 85: 70-85.

- Grimshaw, G.M., Adelstein, A., Bryden, M.P., and MacKinnon, G.E., 1998, "First language acquisition in adolescence: Evidence for a critical period for verbal language development", *Brain and Language*, 63: 237-55.
- Haesler, S., K. Wada, *et al*. (2004). "FoxP2 expression in avian vocal learners and non-learners." *Journal of Neuroscience* 24(13): 3164-3175.
- Hakuta, K., Bialystok, E., and Wiley, E. (2003). "Critical evidence: a test of the critical-period hypothesis for second-language acquisition", *Psycholigical Science,* 141: 31-8.
- Hamilton, R.H. and Pascual-Leone, A. (1998). "Cortical plasticity associated with Braille learning", *Trends in Cognitive Science*, 2: 168-174.
- Harman, G.H. (1967). "Psychological Aspects of the Theory of Syntax", *Journal of Philosophy*, 64: 75-87.
- Harman, G. (1969). Comments on Linguistic Competence and Nativism. *Synthese, 19*, 410-424.
- Harnad, S.R., ed. (1987). *Categorical Perception : The Groundwork of Cognition*, Cambridge: Cambridge University Press.
- Harris, Z.S. (1951). *Structural Linguistics*, Chicago: University of Chicago Press.
- Hauser, M.D., Chomsky, N., and Fitch, W.Y. (2002). "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?", *Science,* 298: 1569-1579.
- Hauser, M. D. and McDermott, J. (2003). "The evolution of the music faculty: a comparative perspective", *Nature Neuroscience*, 6: 663-668.
- Hauser, M.D., Weiss, D., and Marcus, G. (2002). "Rule learning by cotton-top tamarins", *Cognition*, 86: 15-22.
- Hirsh-Pasek, K., Trieman, R. and Schneidermann, M. (1984). "Brown and Hanlon Revisited: Mothers' sensitivity to ungrammatical forms", *Journal of Child Language*, 11: 81-88.
- Holland, A.L., Fromm, D.S., DeRuyter, F. and Stein, M. (1996). "Treatment efficacy: aphasia", *Journal of Speech and Hearing Research*, 39: S27-36.
- Hornstein, N. and Lightfoot, D., (eds.) (1981). *Explanation in Linguistics: The Logical Problem of Language Acquisition*, London: Longman.
- Hubel, D.H. and Wiesel, T.N. (1970). "The period of susceptibility to the physiological effects of unilateral eye closure in kittens", *Journal of Physiology*, 206: 419-436.
- Jackendoff, R. (1999). "Possible stages in the evolution of the language capacity," *Trends in Cognitive Sciences*, 3:272-279.
- Jain, S., Osherson, D., Royer, J.S. and Sharma, A. (1999). *Systems that Learn*, 2nd Ed., Cambridge, MA: MIT Press.
- Johnson, J. S., & Newport, E. L. (1989). "Critical period effects in 2nd language-learning - the influence of maturational state on the acquisition of english as a 2nd language." *Cognitive Psychology, 21*(1), 60-99.
- Johnson, J.S. and Newport, E.L. (1991). "Critical Period Effects on universal properties of language: the status of subjacency in the acquisition of a second language", *Cognition*, 39: 215-58.
- Karbe, H., Thiel, A., Weber-Luxemberger, G., Herholz, K., Kessler, J. and Heiss, W. (1998). "Brain Plasticity in Poststroke Aphasia: What Is the Contribution of the Right Hemisphere?", *Brain and Language*, 64: 215-230.
- Karmiloff-Smith, A., Grant, J., Berthoud, I., Davies, M., Howlin, P., and Udwin, O. (1997). "Language and Williams Syndrome: How intact is intact?", *Child Development*, 68(2): 246-62.
- ——, J. H. Brown, *et al*. (2003). "Dethroning the myth: Cognitive dissociations and innate modularity in Williams syndrome." *Developmental Neuropsychology* 23(1-2): 227-242.
- Katz, L.C., Weliky, M., and Crowley, J.C. (2000). "Activity and the development of the visual cortex: new perpsectives", in *The New Cognitive Neurosciences, 2nd Ed.*, M.S. Gazzaniga (ed.), Cambridge: MIT Press, pp. 199-212.
- Kegl J., Senghas A., Coppola M. (1999). "Creation through contact: Sign language emergence and sign language change in Nicaragua." In M. DeGraff (ed), *Comparative Grammatical Change: The Intersection of Language Acquisition, Creole Genesis and Diachronic Syntax*, Cambridge MA: MIT Press, pp.179-237.
- Kellerman, E., van Ijzendorn, J., and Takashima, H. (1999). "Retesting a universal: the Empty Category Principle and learners of (pseudo)Japanese", in *The acquisition of Japanese as a second language*, K. Kanno (ed.), Amsterdam: John Benjamins, pp. 71-87.
- Kellerman, E. and Yoshioka, K. (1999). "Inter- and intra-population consistency: a comment on Kanno (1998)", *Second Language Research*, 15: 101-9.

- Kotsoni E., de Haan, M. and Johnson, M.H. (2001). "Categorical perception of facial expressions by 7-month-old infants", *Perception*, 30 (9): 1115-1125.
- Kuhl, P.K. (1994). "Learning and Representation in Speech and Language", *Current Opinion in Neurobiology,* 4: 812-822.
- —— (2000). "A new view of language acquisition", *PNAS,* 97 (22): 11850-11857.
- ——, and Miller, J.D. (1975). "Speech Perception by the Chinchilla: Voiced-Voiceless Distinction in Alveolar Plosive Consonants", *Science*, 190: 69-72.
- ——, Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U., Lacerda, F. (1997a). "Cross-Language Analysis of Phonetic Units in Language Addressed to Infants", *Science,* 277: 684-686.
- ——, Kiritani, S., Degughi, T., Hayashi, A., Stevens, E.B., Dugger, C.D. and Iverson, P. (1997b). "Effects of language experience on speech perception: American and Japanese infants' perception of /ra/ and /la/", *Journal of the Acoustical Society of America*, 100: 2425-38.
- Kujala, T., Alho, K., Huotilainen, M., Ilmoniemi, R.J., Lehtoski, A., Leionen, A., Rinne, T., Salonen, O., Sinkkonen, J., Standertskjold-Nordenstam, C.G., and Naatanen, R. (1997). "Electrophysiological evidence for cross-modal plasticity in humans with early- and late-onset blindness", *Psychophysiology*, 34: 213-6.
- Jones, P. (1995). Contradictions and unanswered questions in the Genie case: A fresh look at the linguistic evidence. *Language and Communication, 15*(3), 261-280.
- Lai CSL, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001). "A forkhead-domain gene is mutated in a severe speech and language disorder", *Nature*, 413 (6855): 519-523.
- ——, D. Gerrelli, *et al*. (2003). "FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder." *Brain* 126: 2455-2462.
- Lasnik, H. (1989). "On Certain Substitutes for Negative Data", Demopoulos and Marras (1989:89-105).
- Lasnik, H. and Uriagereka, J. (1986). *A Course in GB Syntax: Lectures on Binding and Empty Categories* (Current Studies in Linguistics), Cambridge: MIT Press.
- Lenneberg, E.H. (1964). "A biological perspective of language", in E.H. Lenneberg, ed., *New Directions in the Study of Language*, Cambridge: MIT Press. Pp.65-88
- —— (1967). *Biological Foundations of Language,* New York: Wiley.
- Levy, Y. and Hermon (2003). "Morphological abilities of Hebrew-speaking adolescents with Williams syndrome," *Developmental Neuropsychology*, 23(1-2):59-83
- Lewontin, R. (1998). "The Evolution of Cognition: Questions we will Never Answer," In R. Sternberg and D. Scarborough (eds.), *An Invitation to Cognitive Science, Vol. 4: Methods, Models and Conceptual Issues*, Cambridge, MIT Press, pp.107-32.
- Lenneberg, E.H. (1964). "A biological perspective of language", in E.H. Lenneberg, ed., *New Directions in the Study of Language*, Cambridge: MIT Press. Pp.65-88
- —— (1967). *Biological Foundations of Language,* New York: Wiley.
- Lieberman, P. (2002). "On the nature and evolution of the neural bases of human language," *American Journal of Physical Anthropology,* 119, pp. 36-62.Liegeois, F., T. Baldeweg, *et al*. (2003). "Language fMRI abnormalities associated with FOXP2 gene mutation." *Nature Neuroscience* 6(11): 1230-1237.
- Lumsden, J.S. (1999). "Language acquisition and creolization", *Language Creation and Language Change: Creolization, Diachrony and Development*, M. DeGraff (ed.), Cambridge: MIT Press, pp.129-57.
- Maratsos, M.P. (1989). "Innateness and plasticity in language acquisition," in M.L. Rice and R.L. Schiefelbusch, eds., *The Teachability of Language*, Baltimore, Paul H. Brooks, 105-25.
- Markson, L. (2006). Core mechanisms of word learning. In *Processes of Change in Brain and Cognitive Development: Attention and Performance Xxi* (pp. 111-128). New York: Oxford Univ Press.
- Martin, R.C. (2003). "Language Processing: Functional Organization and Neuroanatomical Basis", *Annual Review of Psychology*, 54:55-89.
- Mayberry R.I., Eichen, E.B. (1991). "The Long-Lasting Advantage of Learning Sign Language in Childhood - Another Look at the Critical Period for Language-Acquisition", *J Mem Lang,* 30 (4): 486-512.
- Macdonald, C., (ed) (1995). "Tacit Knowledge" in *Philosophy of Psychology: Debates on Psychological Explanation,* Cambridge: Blackwell.
- McGonigle, B., Chalmers, M., and Dickinson, A. (2003). "Concurrent disjoint and reciprocal classification by Cebus apella in seriation tasks: evidence for hierarchical organization", *Animal Cognition* Vol. 6, No. 3: 185-197.

- Maess, B., Koelsch, S., Gunter, T. C. & Friederici, A. D., 2001, "Musical Syntax is processed in Broca's area: an MEG study", *Nature Neuroscience,* 4: 540-545.
- Marcus, G. F. (1998). "Can connectionism save constructivism?" *Cognition*, 66: 153-182.
- —— (2001). *The Algebraic Mind: Integrating Connectionism and Cognitive Science*, Cambridge, MA: MIT Press.
- —— (2004). *The Birth of the Mind: How a Tiny Number of Genes Creates the Complexity of Human Thought*, Basic Books, New York.
- Margolis, E. (1998). "How to Acquire a Concept", *Mind and Language,* 13(3): 347-369.
- Matthews, R.J. (2001). "Cowie's Anti-Nativism", *Mind and Language*, 16: 215-230.
- von Melchner, L., Pallas, S.L. and Sur, M. (2000). "Visual behaviour mediated by retunal projections directed to the auditory pathway", *Nature*, 404: 871-76.
- Mervis, C. B., B. F. Robinson, *et al.* (2000). "The Williams syndrome cognitive profile." *Brain and Cognition,* 44(3): 604-628.
- Mintz T.H. (2002). "Category induction from distributional cues in an artificial language", *Memory and Cognition*, 30: 678-686.
- Moerk, E. (1991). "Positive evidence for negative evidence", *First Language*, 11: 219-51.
- Mohr, J.P., Pessin, M.S., Finkelstein, S., Funkenstein, H.H., Duncan, G.W. and Davis, K.R. (1978). "Broca aphasia: pathological and clinical", *Neurology*, 28: 311-24.
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). "Structural packaging in the input to language-learning - contributions of prosodic and morphological marking of phrases to the acquisition of language." *Cognitive Psychology, 19*(4), 498-550.
- Morgan, J.L. and Travis, L.L. (1989). "Limits on negative information in language input", *Journal of Child Language,* 16: 531-552.
- Morgan, J.L., Bonamo, K.M. and Travis, L.L. (1995). "Negative Evidence on Negative Evidence", *Developmental Psychology*, 31: 180-197.
- Morris, C. A. and C. B. Mervis (2000). "Williams syndrome and related disorders." *Annual Review of Genomics and Human Genetics* 1: 461-484.
- Mufwene, S.S. (1999). "On the language bioprogram hypothesis: hints from Tarzie", *Language Creation and Language Change: Creolization, Diachrony and Development*, M. DeGraff (ed.), Cambridge: MIT Press, pp.95-127.
- Muller R.A., Rothermel R.D., Behen M.E., Muzik O., Chakraborty P.K., and Chugani H.T. (1999). "Language organization in patients with early and late left-hemisphere lesion: a PET study", *Neuropsychologia*, 37: 545-557.
- Myhill, J. (1991). "Typological test analysis: tense and aspecte in creoles and second languages", in *Crosscurrents in Second Language Acquisition and Linguistic Theories*, T. Huebner and C.A. Ferguson (eds.), Amsterdam: John Benjamins, pp. 93-121.
- Nazzi, T. and Ramus, F. (2003). "Perception and acquisition of linguistic rhythm by infants", *Journal of Speech Communication*, 41: 233-43.
- Nelson D.A., Marler P. (1989). "Categorical perception of a natural stimulus continuum: birdsong", *Science*, 244: 976-8.
- Newell, F.N., Bulthoff, H.H. (2002). "Categorical perception of familiar objects", *Cognition* 85 (2): 113-143
- Newmeyer, F.J. (1986). *Linguistic Theory in America*, 2nd Ed., New York: Academic Press.
- —— (1997). *Generative Linguistics: A Historical Perspective*, London and New York: Routledge.
- Newport, E.L. (1990). "Maturational constraints on language learning", *Cognitive Science,* 14: 11-28.
- —— (2001). "Reduced input in the acquisition of signed languages: contributions to the study of creolization", *Language Creation and Language Change: Creolization, Diachrony and Development*, M. DeGraff (ed.), Cambridge: MIT Press, pp.161-78.
- Newport, E., Gleitman, H. and Gleitman, L. (1977). "Mother, please, I'd rather do it myself: some effects and non-effects of maternal speech style", in *Talking to Children: Language Input and Acquisition*, Snow, C. and Ferguson, C. (eds.), New York: Cambridge University Press, pp.109-50.
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*(2), 127-162.

- Nishimura, H., Hashikawa, K., Doi, K., Iwaki, T., Watanabe, Y., Kusuoka, H., Nishimura, T. and Kubo, T. (1999). *Nature,* 397: 116.
- Norbury, C. F., D. V. M. Bishop, and Briscoe, J. (2002). "Does impaired grammatical comprehension provide evidence for an innate grammar module?" *Applied Psycholinguistics* 23(2): 247-268.
- O'Brien E.K., Zhang X.Y., Nishimura C., Tomblin J.B., Murray J.C. (2003). "Association of specific language impairment (SLI) to the region of 7q31", *American Journal Of Human Genetics* 72 (6): 1536-1543
- Odling-Smee, J., Laland, K. and Feldman, M. 1996, "Niche Construction," *American Naturalist*, 147:641-8.
- Pastore, R.E., Li, X.F. and Layer, J.K. (1990). "Categorical perception of nonspeech chirps and bleats", *Percept Psychophys*, 48(2):151-6.
- Pena, M., Bonatti, L.L., Nespor, M., and Mehler, J. (2002). "Signal-driven computations in speech processing", *Science*, 298: 604-7.
- Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*, Cambridge, MA: Bradford/MIT Press.
- —— (1991). "Rules of Language." *Science*, 253:530-35.
- —— (1994). *The Language Instinct*, New York: Harper Collins.
- —— (1997). *How the Mind Works*, New York, Norton.
- —— and Bloom, P. (1990). "Natural Language and Natural Selection," *Behavioral and Brain Sciences*, 13, pp.707-84.
- Posner, M.I. and Raichle, M.E. (1994). *Images of Mind*. 2nd Ed., New York: W H Freeman & Co.
- Prinz, J.J. (2002). *Furnishing the Mind: Concepts and their Perceptual Basis*, Bradford Books/MIT Press.
- Pullum G.K. and Scholz, B.C. (2002). "Empirical assessment of stimulus poverty arguments", *The Linguistic Review,* 19 9-50.(Special issue, nos. 1-2: 'A Review of "The Poverty of Stimulus Argument",' ed. N. Ritter).
- Putnam, H. (1971). "The 'Innateness Hypothesis' and explanatory models in linguistics", in *The Philosophy of Language,* J. Searle (ed.), London: Oxford University Press, pp.130-9.
- Quartz, S.R. and Sejnowski, T.J. (1997). "The Neural Basis of Cognitive Development: A Constructivist Manifesto," *Brain and Behavioral Sciences*, 20:537-596.
- Ramus, F., Hauser, M.D., Miller, C., Morris, D., Mehler, J. (2000). "Language Discrimination by Human Newborns and by Cotton-Top Tamarin Monkeys", *Science,* 288: 349-351.
- Redington, M. and Chater, N. (1998). "Connectionist and statistical approaches to language acquisition: A distributional perspective", *Language and Cognitive Processes*, 13: 129-191.
- Rice, M. L. and K. Wexler (1996). "Toward tense as a clinical marker of specific language impairment in English-speaking children." *Journal of Speech and Hearing Research* 39(6): 1239-1257.
- Ritter, N.A. (2002). 'A Review of "The Poverty of Stimulus Argument",' *The Linguistic Review* (Special Issue), 19:1-2.
- Roeper, T. and Williams, E., (eds.) (1987). *Parameter Setting: Studies in Theoretical Psycholinguistics*, Vol. 4, Dordrecht: Kluwer.
- Sadato, N. Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M.P., Dold, G. and Hallett, M. (1996). "Activation of the primary visual cortex by Braille reading in blind subjects", *Nature*, 380: 526-8.
- Saffran, E.M. (2000). "Aphasias and the relationship of language and brain", *Seminars in Neurology,* 20: 409-418.
- Saffran, J.R., Aslin, R.N. and Newport, E.L. (1996). "Statistical Learning by 8-Month-Old Infants", *Science*, 274: 1926-28.
- Saffran, J.R. (2002). "Constraints on statistical language learning", *Journal of Memory and Language*, 47: 172-96.
- Saffran, J.R. and Thiessen, E.D. (2003). "Pattern induction by infant language learners", *Developmental Psychology*, 40: 484-94.
- Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy, 4*(2), 273-284.
- Samet, J. (1986). "Troubles with Fodor's Nativism", *Midwest Studies in Philosophy*, X. Boston: Blackwells. 575-594.

- Sampson, G. (1989). "Language Acquisition: Growth or Learning?", *Philosophical Papers*, XVIII: 203-40.
- —— (2002). "Exploring the richness of the stimulus", *The Linguistic Review*, 19: 73-104. (Special issue, nos. 1-2: 'A Review of "The Poverty of Stimulus Argument"', N. Ritter (ed.)).
- Sapir, E. (1921). *Language: An Introduction to the Study of Speech*, Hillsdale, NJ: Erlbaum.
- Saxton M. (1997). "The contrast theory of negative input", *Journal of Child Language*, 24:139-61.
- ——, Kulcsar B., Marshall G., and Rupra M.J., 1988, "Longer-term effects of corrective input: an experimental approach", *Child Language*, 25: 701-21.
- Scharff, C. and S. A. White (2004). Genetic components of vocal learning. *Behavioral Neurobiology of Birdsong*. 1016: 325-347.
- Scholz, B.C. and Pullum, G.K. (2002). "Searching for arguments to support linguistic nativism", *The Linguistic Review*, 19: 185-224. (Special issue, nos. 1-2: 'A Review of "The Poverty of Stimulus Argument",' N. Ritter(ed.))
- Schulte, O. (2006). "Formal Learning Theory", The Stanford Encyclopedia of Philosophy (Spring 2006 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2006/entries/learning-formal/>.
- Sharwood Smith, M. (1994). *Second Language Learning: Theoretical Foundations*, London and New York: Longman.
- Shimojo, S. and Shams, L. (2001). "Sensory modalities are not separate modalities: plasticity and interactions", *Current Opinion in Neurobiology*, 11: 505-9.
- Shinohara, T. (1994). "Rich classes inferable from positive data: length-bounded elementary formal systems", *Information and Computation*, 108: 175-186.
- Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology, 49*(4), 370-407.
- Skinner, B.F. (1957). *Verbal Behavior*, NY: Prentice Hall.
- Skuse, D.H. (1993). "Extreme deprivation in early childhood", in *Language development in exceptional circumstances*, D. Bishop and K. Mogford (eds.), Hillsdale, NJ: Erlbaum, pp. 29-46.
- Skyrms, B. (1996). *The Evolution of the Social Contract*, Cambridge, Cambridge University Press.
- —— (2002). *The Stag Hunt and the Evolution of Social Structure*, Cambridge, Cambridge University Press.
- Smith, E.E. and Medin, D.L. (1981). *Categories and Concepts*, Cambridge, MA: Harvard University Press.
- Smolensky, P. (1991). "Connectionism, Constituency and the Language of Thought", in *Meaning in Mind: Fodor and His Critics*, B. Loewer and G. Rey, (eds.), Cambridge MA: Blackwells, 201-227.
- Soames, S. (1984). "Linguistics and Psychology", *Linguistics and Philosophy*, 7: 155-79.
- Sokolov, J.L. and Snow, C.E. (1991). "The Premature Retreat to Nativism," *Behavioral and Brain Sciences,* 14:635-6.
- Sterelny, K. (1989). "Fodor's Nativism", *Philosophical Studies*, 55: 119-141.
- —— 2003, *Thought in a Hostile World: The Evolution of Human Cognition*, London, Blackwells.
- Stich, S.P. (1971). "What every speaker knows", *Philosophical Review*, 80: 476-96.
- —— (1978). "Beliefs and Subdoxastic States," *Philosophy of Science*, 45, pp.499-518.
- Stiles, J. (2000). "Neural plasticity and cognitive development", *Developmental Neuroscience*, 18 (2): 237-272.
- Stromswold, K. (2000). "The cognitive neuroscience of language acquisition", in *The New Cognitive Neurosciences, 2nd Ed* M.S. Gazzaniga (ed.), Cambridge: MIT Press, pp.269-80.
- Takahashi K., Liu, F.C., Hirokawa, K., and Takahashi, H. (2003). "Expression of Foxp2, a gene involved in speech and language, in the developing and adult striatum", *Journal of Neuroscience Research*, 73 (1): 61-72.
- Tallal, P. (1980). "Auditory Temporal Perception, Phonics, And Reading Disabilities In Children." *Brain And Language* 9(2): 182-198.
- Tallal, P., R. E. Stark, *et al.* (1985). "Identification Of Language-Impaired Children On The Basis Of Rapid Perception And Production Skills." *Brain and Language* 25(2): 314-322.
- Teramitsu, I., L. C. Kudo, *et al.* (2004). "Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction." *Journal of Neuroscience* 24(13): 3152-3163.

- Thiessen, E.D. and Saffran, J.R. (2003). "When cues collide: Use of stress and statistical cues to word boundaries by 7-to 9-month-old infants", *Developmental Psychology*, 39: 706-16.
- Thomas, M. and Karmiloff-Smith, A. (2002): "Are developmental disorders like cases of adult brain damage? Implications from connectionist modeling." *Behavioral and Brain Sciences*, 25: 727—788
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*, Cambridge, MA, Harvard University Press.
- —— (2000). "Two Hypotheses about Primate Cognition," in C. Heyes and L. Huber (Eds.), *Evolution of Cognition*, Cambridge, MIT Press, pp.165-84.
- —— (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.
- vanderLely, H. K. J. and L. Stollwerck (1997). "Binding theory and grammatical specific language impairment in children." *Cognition* 62(3): 245-290.
- van der Lely, H. K. J. and V. Christian (2000). "Lexical word formation in children with grammatical SLI: a grammar-specific versus an input-processing deficit?" *Cognition* 75(1): 33-63.
- van der Lely, H. K. J. and M. T. Ullman (2001). "Past tense morphology in specifically language impaired and normally developing children." *Language and Cognitive Processes* 16(2-3): 177-217.
- Vargha-Khadem, F., Carr, L.J., Isaacs, E., Brett, E., Adams, C. and Mishkin, M. (1997). "Onset of speech after left hempisherectomy in a nine-year-old boy", *Brain*, 120: 159-82.
- ——, Watkins, K., Alcock, K., Fletcher, P. and Passingham, R., 1995, "Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder," *Proc. Natl. Acad. Sci. USA*, 92:930-33.
- Volterra, V., O. Capirci, *et al.* (1996). "Linguistic abilities in Italian children with Williams syndrome." *Cortex* 32(4): 663-677.
- Watkins, Dronkers & Vargha-Khadem, 'Behavioral analysis of an inherited speech and language disorder: comparison with acquired aphasia.' *Brain* (2002), 125, 452-464.
- White, L. (2003). *Second Language Acquisition and Universal Grammar*, Cambridge Textbooks in Linguistics, Cambridge: Cambridge University Press.
- Wimsatt, W.C. (1999). "Generativity, Entrenchment, Evolution and Innateness: Philosophy, evolutionary biology, and conceptual foundations of science", in *Where Biology Meets Psychology: Philosophical Essays*, V. Hardcastle (ed.), Cambridge: Bradford Books/MIT Press, pp.139-179.
- Wright, A.A, Rivera, J.J., Hulse, S.H., Shyan, M. and Neiworth, J.J. (2000). "Music perception and octave generalization in rhesus monkeys", *Journal of Experimental Psychology: General*, 129 (3): 291-307.
- van Gelder, T. (1995). "What Might Cognition Be, If Not Computation?", *Journal of Philosophy*, 92: 345-381.
- Vargha-Khadem F., Watkins, K.E., Price, C.J., Ashburner, J., Alcock, K.J., Connelly, A., Frackowiak, R.S., Friston, K.J., Pembrey, M.E., Mishkin, M., Gadian, D.G., and Passingham, R.E. (1998). "Neural basis of an inherited speech and language disorder", *Proceedings of the National Academy of Sciences, USA*, 95 (21): 12695-12700.
- Watkins, K.E., Dronkers, N.F., and Vargha-Khadem, F. (2002). "Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia", *Brain*, 125: 452-464.
- Zhang J.Z., Webb, D.M., and Podlaha, O. (2002). "Accelerated protein evolution and origins of human-specific features: FOXP2 as an example", *Genetics,* 162 (4): 1825-1835.

# Academic Tools

- How to cite this entry.
- Preview the PDF version of this entry at the Friends of the SEP Society.
- Look up this entry topic at the Indiana Philosophy Ontology Project (InPhO).
- Enhanced bibliography for this entry at PhilPapers, with links to its database.

# Other Internet Resources

- [COLT: Computational Learning Theory](#)

# Related Entries

## Acknowledgments

Stanford | Center for the Study of Language and Information