

Is the Human Ability to Acquire Languages Innate?

Written for UCL's Philosophy of Language module. I argue for linguistic nativism: children converge on the same grammar despite input that underdetermines it, which is the core of Chomsky's Poverty of the Stimulus argument. I work through the standard behaviourist objection, then a more current one: that LLMs seem to pick up grammatical competence from exposure alone, no innate constraints required. I argue that while nativism survives both, the LLM objection does force a weaker, more input-sensitive version of the theory than the one I started with.

ARGUMENT IN BRIEF

1. The claim is about natural languages, acquired in a critical period, largely without explicit instruction, and distinct from artificial ones.
2. Poverty of the Stimulus: the input a child receives underdetermines the grammar they end up with, yet children converge on the same grammar anyway.
3. The best explanation is innate internal constraints on the hypotheses a learner will entertain, which is Universal Grammar.
4. Objection one, behaviourism: acquisition is conditioning on input. This does not account for the convergence, and I reject it.
5. Objection two, LLMs: models appear to acquire grammatical competence from exposure alone, with no innate grammar.
6. Nativism survives, but not intact. What LLMs show is that input does far more work than the strong version of the argument allows, so the conclusion has to be weakened to constraints plus substantial input, not constraints alone.

This essay argues that language acquisition is innate in humans. I distinguish between two types of language: natural and artificial. After, I focus on the linguistic nativist claim that children have innate internal constraints that assist them in acquiring natural languages. I outline Noam Chomsky's Poverty of the Stimulus argument and the concept of Universal

Grammar to support this position. I then consider two objections to nativism before ultimately concluding that, although human language acquisition ultimately relies on internal constraints and is thus innate, external linguistic input can still play a role in facilitating it.

Natural languages are best understood as those that develop organically over time, such as English and Mandarin. They are typically characterised by the critical biological period during which an individual begins to learn them, the guidance required, and the degree of conscious access to them. Most individuals (if said natural language is one's first language) pick up a natural language from birth to around the ages of six or seven. Children learn these languages very quickly during this period. Additionally, no explicit didactic instruction is often required. Children may receive formal instruction in how these languages work, but most of their ability to learn them develops passively. Furthermore, access to one's natural language competence is often inaccessible. If one were to ask why I know how the English language works, I would struggle to point to or articulate any formal grammatical rules which I had learnt. Rather, my confidence in the English language is largely subconscious. For example, take the following sentences:

- (i) The idea was famously Italian.
- (ii) It true that the idea had attributed to an Italian?

English speakers can intuitively claim that while (i) makes grammatical sense, (ii) does not. However, if asked why, most English speakers would struggle to articulate that (ii) does not obey rules surrounding auxiliary verbs and clause structures.

By contrast, artificial languages are dedicated language tools, limited to describing specific scenarios and used only in certain contexts. Examples include chess notation or first-order propositional logic. Additionally, there is no designated critical period for when these languages are typically learnt. Rather, an individual can learn them at any stage in life. Furthermore, these languages require explicit instruction and cannot be acquired through exposure alone. First-order logic requires formal teaching (e.g., in a university degree program) to be learnt. Finally, these languages reveal a conscious insight into the speaker's mind. The learner has introspective access to the language's grammar that can be articulated. A student of first-order language can point to specific syntactical and grammatical rules to explain why a certain proposition is flawed while another is not.

Having explained the natural-artificial language distinction, I shall outline the nativist position. The nativist argues that humans have innate capacities to acquire *natural languages*. This explains how children acquire such mastery of a language despite imperfect linguistic input. An argument for this position is Chomsky's Poverty of the Stimulus (PoS) argument (Davis and Gillon 15–16):

- P1. Children acquire mastery over a specific grammar ($\hat{A}$)
- P2. The exposure (the stimulus) is often highly complex in its acoustic signal and can be compatible with other grammars ($A_1 \dots A_n$)
- P3. Nothing in the input they receive necessarily rules out other grammars
- P4. Children do not always receive sufficient evidence to eliminate incorrect grammars
- P5. Children nonetheless converge on the same grammar ($\hat{A}$)
- C. Thus, language acquisition cannot be explained by input alone; something innate must guide it.

This can be explained through the following sentences:

- (iii) The somnambulating man is tired.
- (iv) Is the somnambulating man tired?

To transform declarative A into interrogative B, the speaker must take the auxiliary of the main clause and move it (RULE 1). Additionally, consider the following:

- (v) The man who is Belgian is pandiculating.
- (vi) Is the man who Belgian pandiculating?
- (vii) Is the man who is Belgian pandiculating?

For these sentences, a similar thing has occurred as in (iii) and (iv), except there are now two verbs. RULE 1 remains compatible with the child's earlier input but yields an incorrect result in cases such as (vi). The movement of the two verbs follows a second rule that is more complicated, as it involves more steps (RULE 2). However, when learning English, children are not explicitly taught clauses and these rules. Additionally, they do not always receive correction for errors such as those in (vi). Despite this, children are aware that one applies to (iii) and (iv), and a different one applies to (v), (vi) and (vii). Children converge on this grammatical nuance, indicating that linguistic stimuli alone do not determine the necessary grammar. Thus, linguistic input alone cannot account for language acquisition, and some intuitive constraints must guide the hypotheses available to the learner.

Having clarified the PoS argument, the nativist argues that linguistic inputs alone cannot account for the grammar that children converge on, since it is compatible with many grammatical hypotheses. Thus, the PoS argument leads the nativist to show how children converge on the same grammar(s). This is done through an account of Universal Grammar.

Universal Grammar refers to the innate human endowment that accounts for grammatical and linguistic acquisition (Davis and Gillon 15). It is best thought of as a system of linguistic constraints on the grammars that a child can acquire. Universal grammar is not specific to any particular natural language. Rather, it is a biological phenomenon that allows humans to acquire any natural language. This point is best seen when a sentence is semantically ambiguous or uninterpretable but syntactically well-formed:

(viii) Milquetoast frames surrender relentlessly.

If sentences and their grammar were learnt solely by semantic associations, then (viii) would be deemed ungrammatical. However, speakers can judge that (viii) makes grammatical sense despite its semantic oddity, suggesting that grammatical competence lies in something beyond surface-level interpretations of sentences and their meanings. Speakers are confident in their ability to assess the grammatical value of sentences they have (likely) never encountered before. This additionally shows that linguistic understanding cannot rely solely on memorised usage. Linguistic understanding seems, therefore, to require an algorithmic and compositional system that designates structure independently of meaning (Chierchia and McConnell-Ginet 6).

Having laid out the nativist position, I shall address potential objections. The first is a behaviourist objection that denies the appeal to any internal mental states and internal conception of grammar in language acquisition (Davis and Gillon 9). Behaviourists argue that language acquisition is explained by linguistic inputs and stimulus responses, reinforced by general learning patterns which may include imitation, pattern recognition, etc. Language acquisition is treated as a skill that is learnt in much the same way as other skills; a child learns to ride a bike through continued practice and pattern recognition of what works and what does not. Grammatical knowledge is grounded in experience rather than in internal, intuitive guidance, so there is no need to account for language constraints, as suggested by the nativist. Rather, any errors in a child's ability to learn a language are corrected by gradual exposure and feedback. This clearly opposes the nativist position that language acquisition requires innate mental states.

In response, the nativist may raise two points: overgeneration and the decisions surrounding rule choice. Firstly, children often construct their own grammatical rules when learning a language in ways that go beyond any linguistic input they have received. Take the following sentences:

(ix) She feels miserable

(x) She feeled miserable

Here, one can imagine a child saying (x) in an attempt to convey the same meaning as in (ix), but in the past tense. Of course, (x) is grammatically incorrect. However, a child will have never heard an adult say "feeled", so this word will never have been part of their linguistic input. This error is therefore not an imitation of anything the child has heard, nor does their environment reinforce it. Rather, it has come from the internal rule the child intuitively has, of adding '-ed' to the end of verbs to form the past tense. This shows how language acquisition does not consist merely in the memorisation of external stimuli. Rather, the child must be constructing an internal grammatical system that is not merely based on habit or input. Behaviourism rejects this but cannot account for the internal, rule-governed grammatical system that children possess.

Additionally, the nativist may respond by referring to the decision a child makes when choosing language rules. This reinforces the claim that the rules children adopt are constrained. This is seen in (iii), (iv) and (v), (vi), (viii). Whereas one rule governing the movement of auxiliary verbs is used to convert (iii) into the interrogative (iv), a different and more complex rule is used to turn (v) into the interrogative (viii). Both rules are compatible with the linguistic inputs that children have access to, yet children do not adopt the simpler rule (applying to (iii) and (iv)) when dealing with instances of (viii). Yet, if behaviourism were true, this would not be the case; if children were merely relying on pattern recognition and selecting the simplest pattern, they would always choose the first rule. However, of the two hypotheses (rules) available to the child, children will always converge on the second, more complex one for (viii). This indicates that children do not consider applying RULE 1, despite its compatibility with their linguistic input and never having been explicitly told to do so. Thus, there must be internal constraints that dictate which hypothesis children apply in different scenarios. Behaviourism does not explain why one hypothesis is selected over another.

Taken together, these two replies reject the behaviourist objection. The nativist account shows how children construct their own grammatical rules and thus do not rely solely on external linguistic inputs. Additionally, the rules to which children have 'access' are restrained by internal constraints. Behaviourism appeals to pattern recognition, external stimuli and memorisation and is thus unable to explain why children converge on the same grammatical rules and linguistic competence.

Having rejected the behaviourist objection, nativism remains open to further criticism. A strong criticism argues that the development of LLMs (large language models) suggests that language acquisition does not require internal constraints. LLMs, such as ChatGPT, can construct sentences in a language with grammatical 'competence' despite having no innate rules and being trained solely on language data. Rather, they are fed vast amounts of data from which they recognise patterns in language and text.

This objection targets the PoS argument, which argues that linguistic input alone is insufficient to account for how learners converge on the same universal grammar. However, if LLMs can acquire language from exposure alone, then the claim that grammar cannot be learned from simply linguistic inputs is flawed. Importantly, the relevance of the LLM objection does not lie in explaining any similarities between artificial language models and humans, but in placing pressure on the nativist's claim that grammatical structure cannot be learned from linguistic input alone. If correct, this objection shows that the PoS argument may underestimate the ability of linguistic stimuli to motivate language acquisition.

The nativist may reply to this objection by highlighting the difference between LLMs and humans. Unlike LLMs, children receive fewer 'data points' during the critical period of learning a natural language, and these data are of varying quality. While an LLM is fed data

optimised to maximise its performance, children are exposed to acoustic signals that give rise to different interpretations of the same expression (e.g. hearing "The man saw the boy with a telescope" yields different interpretations of its meaning (Davis and Gillon 15)). Thus, while LLMs receive many accurate data inputs over an indefinite period, children are exposed to a far more limited set of ambiguous acoustic signals over a much shorter period (the critical period). This is significant because the PoS argument concerns *human* language acquisition. This objection does not necessarily undermine the linguistic nativist's claims. However, the nativist must still explain how it is that LLMs deny the presence of an innate linguistic structure.

The nativist may respond to the claim that if LLMs demonstrate grammatical competence, then language acquisition is not innate. Even if the LLM objection does not claim that LLMs possess human-like grammatical competence, it argues that exposure to linguistic input alone makes language acquisition possible. However, the nativist argues that LLMs lack the human cognitive capacity for language acquisition. Although LLMs may exhibit language-like behaviour, this does not necessarily mean they possess the capacity to acquire language. This is done by appealing to the original criteria that characterise natural language, namely, that its acquisition is a biological process that unfolds over time. By contrast, LLMs do not converge on a grammar over time but reach their grammatical 'competence' by optimising outputs through large data inputs. Thus, the objection does not show that human language acquisition is not grounded in internal constraints.

Furthermore, the nativist may argue that LLMs possess internal structure that constrains their language hypotheses. LLMs comply with various safety constraints that dictate how information is presented and what outputs are possible. These work with pre-built objectives to optimise response efficiency. This supports the claim that language acquisition requires prior (innate) constraints. If this were not the case, the learner would have access to many language hypotheses compatible with their linguistic inputs. Without constraints, the learner would be unable to choose one set of hypotheses over another. Of course, the constraints that an LLM has are not directly equivalent to those of a child, yet they perform the same function in guiding the learner and constraining the set of available hypotheses.

However, this objection may force the nativist to modify their position. The existence of LLMs may show that the PoS argument derives too strong a conclusion. LLMs show that linguistic inputs can lead to language-like performance (even when aided by internal constraints); extensive linguistic inputs, when supported by constrained systems, give rise to sophisticated, language-like competence. This does not reject claims of innateness, but suggests a different account. The nativist defends the central claim that internal constraints limit the language hypotheses available to the learner, but now acknowledges the significant role of external linguistic input in shaping language acquisition. Thus, in human language

acquisition, even low-quality linguistic input plays a role. Their innate constraints interact with their linguistic input to produce an internal grammar system that enables children to acquire language.

In conclusion, I have argued that the human capacity to acquire language is innate but not independent of linguistic inputs. The PoS argument shows that linguistic input cannot be the only factor enabling a child to acquire grammar, thus supporting the position that there must be innate constraints. While the nativist can reject behaviourist objections, the LLM objection reveals that internal constraints are required for acquiring grammar and language. However, the LLM objection also shows the significant role that external linguistic inputs play in one's acquisition of language. They help shape the overall grammar a learner acquires and work with internal constraints to determine the grammar a learner acquires as part of their linguistic understanding.

WORKS CITED

Chierchia, Gennaro, and Sally McConnell-Ginet. *Meaning and Grammar, Second Edition: An Introduction to Semantics*. Penguin Random House LLC, 2000.

Davis, Steven, and Brendan S. Gillon, editors. *Semantics: A Reader*. Oxford University Press, USA, 2004.

RELATED

[Multiple Realizability](#)

[Enoda, and why I keep AI out of it](#)