

# Evidence Rebutts Chomsky's Theory of Language Learning

Much of Noam Chomsky's revolution in linguistics — including its account of the way we learn languages — is being overturned

By Paul Ibbotson & Michael Tomasello

*Scientific American, November 2016*

[http://www.scientificamerican.com/article/evidence-rebutts-chomsky-s-theory-of-language-learning/?WT.mc\\_id=SA\\_MB\\_20160907](http://www.scientificamerican.com/article/evidence-rebutts-chomsky-s-theory-of-language-learning/?WT.mc_id=SA_MB_20160907)



Picture Credit: OWEN GILDERSLEEVE

The idea that we have brains hardwired with a mental template for learning grammar — famously espoused by Noam Chomsky of the Massachusetts Institute of Technology — has dominated linguistics for almost half a century. Recently, though, cognitive scientists and linguists have abandoned Chomsky's "universal grammar" theory in droves because of new research examining many different languages — and the way young children learn to understand and speak the tongues of their communities. That work fails to support Chomsky's assertions.

The research suggests a radically different view, in which learning of a child's first language does not rely on an innate grammar module. Instead the new research shows that young children use various types of thinking that may not be specific to language at all — such as the ability to classify the world into categories (people or objects, for instance) and to understand the relations among things. These capabilities, coupled with a unique human ability to grasp what others intend to communicate, allow language to happen. The new findings indicate that if researchers truly want to understand how children, and others, learn languages, they need to look outside of Chomsky's theory for guidance.

This conclusion is important because the study of language plays a central role in diverse disciplines — from poetry to artificial intelligence to linguistics itself; misguided methods lead to questionable results. Further, language is used by humans in ways no animal can match; if you understand what language is, you comprehend a little bit more about human nature.

Chomsky's first version of his theory, put forward in the mid-20th century, meshed with two emerging trends in Western intellectual life. First, he posited that the languages people use to communicate in everyday life behaved like mathematically based languages of the newly emerging field of computer science. His research looked for the underlying computational structure of language and proposed a set of procedures that would create "well-formed" sentences. The revolutionary idea was that a computerlike program could produce sentences real people thought were grammatical. That program could also

purportedly explain as well the way people generated their sentences. This way of talking about language resonated with many scholars eager to embrace a computational approach to ... well ... everything.

As Chomsky was developing his computational theories, he was simultaneously proposing that they were rooted in human biology. In the second half of the 20th century, it was becoming ever clearer that our unique evolutionary history was responsible for many aspects of our unique human psychology, and so the theory resonated on that level as well. His universal grammar was put forward as an innate component of the human mind – and it promised to reveal the deep biological underpinnings of the world's 6,000-plus human languages. The most powerful, not to mention the most beautiful, theories in science reveal hidden unity underneath surface diversity, and so this theory held immediate appeal.

But evidence has overtaken Chomsky's theory, which has been inching toward a slow death for years. It is dying so slowly because, as physicist Max Planck once noted, older scholars tend to hang on to the old ways: "Science progresses one funeral at a time."

## In the Beginning

The earliest incarnations of universal grammar in the 1960s took the underlying structure of "standard average European" languages as their starting point – the ones spoken by most of the linguists working on them. Thus, the universal grammar program operated on chunks of language, such as noun phrases ("The nice dogs") and verb phrases ("like cats").

Fairly soon, however, linguistic comparisons among multiple languages began rolling in that did not fit with this neat schema. Some native Australian languages, such as Warlpiri, had grammatical elements scattered all over the sentence – noun and verb phrases that were not "neatly packaged" so that they could be plugged into Chomsky's universal grammar – and some sentences had no verb phrase at all.

These so-called outliers were difficult to reconcile with the universal grammar that was built on examples from European languages. Other exceptions to Chomsky's theory came from the study of "ergative" languages, such as Basque or Urdu, in which the way a sentence subject is used is very different from that in many European languages, again challenging the idea of a universal grammar.

These findings, along with theoretical linguistic work, led Chomsky and his followers to a wholesale revision of the notion of universal grammar during the 1980s. The new version of the theory, called principles and parameters, replaced a single universal grammar for all the world's languages with a set of "universal" principles governing the structure of language. These principles manifested themselves differently in each language. An analogy might be that we are all born with a basic set of tastes (sweet, sour, bitter, salty and umami) that interact with culture, history and geography to produce the present-day variations in world cuisine. The principles and parameters were a linguistic analogy to tastes. They interacted with culture (whether a child was learning Japanese or English) to produce today's variation in languages as well as defined the set of human languages that were possible.

Languages such as Spanish form fully grammatical sentences without the need for separate subjects – for example, *Tengo zapatos* ("I have shoes"), in which the person who has the shoes, "I," is indicated not by a separate word but by the "o" ending at the end of the verb. Chomsky contended that as soon as children encountered a few sentences of this type, their brains would set a switch to "on," indicating that the sentence subject should be dropped. Then they would know that they could drop the subject in all their sentences.

The "subject-drop" parameter supposedly also determined other structural features of the language. This notion of universal principles fits many European languages reasonably well. But data from non-European languages turned out not to fit the revised version of Chomsky's theory. Indeed, the research that had attempted to identify parameters, such as the subject-drop, ultimately led to the abandonment of the second incarnation of universal grammar because of its failure to stand up to scrutiny.

More recently, in a famous paper published in *Science* in 2002, Chomsky and his co-authors described a universal grammar that included only one feature, called computational recursion (although many advocates of universal grammar still prefer to assume there are many universal principles and parameters). This new shift permitted a limited number of words and rules to be combined to make an unlimited number of sentences.

The endless possibilities exist because of the way recursion embeds a phrase *within* another phrase of the same type. For example, English can embed phrases to the right (“John hopes Mary knows Peter is lying”) or embed centrally (“The dog that the cat that the boy saw chased barked”). In theory, it is possible to go on embedding these phrases infinitely. In practice, understanding starts to break down when the phrases are stacked on top of one another as in these examples. Chomsky thought this breakdown was not directly related to language per se. Rather it was a limitation of human memory. More important, Chomsky proposed that this recursive ability is what sets language apart from other types of thinking such as categorization and perceiving the relations among things. He also proposed recently this ability arose from a single genetic mutation that occurred between 100,000 and 50,000 years ago.

As before, when linguists actually went looking at the variation in languages across the world, they found counterexamples to the claim that this type of recursion was an essential property of language. Some languages — the Amazonian Pirahã, for instance — seem to get by without Chomskyan recursion.

As with all linguistic theories, Chomsky’s universal grammar tries to perform a balancing act. The theory has to be simple enough to be worth having. That is, it must predict some things that are not in the theory itself (otherwise it is just a list of facts). But neither can the theory be so simple that it cannot explain things it should. Take Chomsky’s idea that sentences in all the world’s languages have a “subject.” The problem is the concept of a subject is more like a “family resemblance” of features than a neat category. About 30 different grammatical features define the characteristics of a subject. Any one language will have only a subset of these features — and the subsets often do not overlap with those of other languages.

Chomsky tried to define the components of the essential tool kit of language — the kinds of mental machinery that allow human language to happen. Where counterexamples have been found, some Chomsky defenders have responded that just because a language lacks a certain tool — recursion, for example — does not mean that it is not in the tool kit. In the same way, just because a culture lacks salt to season food does not mean salty is not in its basic taste repertoire. Unfortunately, this line of reasoning makes Chomsky’s proposals difficult to test in practice, and in places they verge on the unfalsifiable.

## Death Knells

A key flaw in Chomsky’s theories is that when applied to language learning, they stipulate that young children come equipped with the capacity to form sentences using abstract grammatical rules. (The precise ones depend on which version of the theory is invoked.) Yet much research now shows that language acquisition does not take place this way. Rather young children begin by learning simple grammatical patterns; then, gradually, they intuit the rules behind them bit by bit.

Thus, young children initially speak with only concrete and simple grammatical constructions based on specific patterns of words: “Where’s the X?”; “I wanna X”; “More X”; “It’s an X”; “I’m X-ing it”; “Put X here”; “Mommy’s X-ing it”; “Let’s X it”; “Throw X”; “X gone”; “Mommy X”; “I Xed it”; “Sit on the X”; “Open X”; “X here”; “There’s an X”; “X broken.” Later, children combine these early patterns into more complex ones, such as “Where’s the X that Mommy Xed?”

Many proponents of universal grammar accept this characterization of children’s early grammatical development. But then they assume that when more complex constructions emerge, this new stage reflects the maturing of a cognitive capacity that uses universal grammar and its abstract grammatical categories and principles.

For example, most universal grammar approaches postulate that a child forms a question by following a set of rules based on grammatical categories such as “What (object) did (auxiliary) you (subject) lose (verb)?” Answer: “I (subject) lost (verb) something (object).” If this postulate is correct, then at a given developmental period children should make similar errors across all wh-question sentences alike. But children’s errors do not fit this prediction. Many of them early in development make errors such as “Why he can’t come?” but at the same time as they make this error — failing to put the “can’t” before the “he” — they correctly form other questions with other “wh-words” and auxiliary verbs, such as the sentence “What does he want?”

Experimental studies confirm that children produce correct question sentences most often with particular wh-words and auxiliary verbs (often those with which they have most experience, such as “What does ...”), while continuing to make errors with question sentences containing other (often less frequent) combinations of wh-words and auxiliary verbs: “Why he can’t come?”

The main response of universal grammarians to such findings is that children have the competence with grammar but that other factors can impede their performance and thus both hide the true nature of their grammar and get in the way of studying the “pure” grammar posited by Chomsky’s linguistics. Among the factors that mask the underlying grammar, they say, include immature memory, attention and social capacities.

Yet the Chomskyan interpretation of the children’s behavior is not the only possibility. Memory, attention and social abilities may not mask the true status of grammar; rather they may well be integral to building a language in the first place. For example, a recent study co-authored by one of us (Ibbotson) showed that children’s ability to produce a correct irregular past tense verb – such as “Every day I fly, yesterday I flew” (not “flyed”) – was associated with their ability to inhibit a tempting response that was unrelated to grammar. (For example, to say the word “moon” while looking at a picture of the sun.) Rather than memory, mental analogies, attention and reasoning about social situations getting in the way of children expressing the pure grammar of Chomskyan linguistics, those mental faculties may explain why language develops as it does.

As with the retreat from the cross-linguistic data and the tool-kit argument, the idea of performance masking competence is also pretty much unfalsifiable. Retreats to this type of claim are common in declining scientific paradigms that lack a strong empirical base – consider, for instance, Freudian psychology and Marxist interpretations of history.

Even beyond these empirical challenges to universal grammar, psycholinguists who work with children have difficulty conceiving theoretically of a process in which children start with the same algebraic grammatical rules for all languages and then proceed to figure out how a particular language – whether English or Swahili – connects with that rule scheme. Linguists call this conundrum the linking problem, and a rare systematic attempt to solve it in the context of universal grammar was made by Harvard University psychologist Steven Pinker for sentence subjects. Pinker’s account, however, turned out not to agree with data from child development studies or to be applicable to grammatical categories other than subjects. And so the linking problem – which should be the central problem in applying universal grammar to language learning – has never been solved or even seriously confronted.

## **An Alternative View**

All of this leads ineluctably to the view that the notion of universal grammar is plain wrong. Of course, scientists never give up on their favorite theory, even in the face of contradictory evidence, until a reasonable alternative appears. Such an alternative, called usage-based linguistics, has now arrived. The theory, which takes a number of forms, proposes that grammatical structure is not innate. Instead grammar is the product of history (the processes that shape how languages are passed from one generation to the next) and human psychology (the set of social and cognitive capacities that allow generations to learn a language in the first place). More important, this theory proposes that language recruits brain systems that may not have evolved specifically for that purpose and so is a different idea to Chomsky’s single-gene mutation for recursion.

In the new usage-based approach (which includes ideas from functional linguistics, cognitive linguistics and construction grammar), children are not born with a universal, dedicated tool for learning grammar. Instead they inherit the mental equivalent of a Swiss Army knife: a set of general-purpose tools – such as categorization, the reading of communicative intentions, and analogy making, with which children build grammatical categories and rules from the language they hear around them.

For instance, English-speaking children understand “The cat ate the rabbit,” and by analogy they also understand “The goat tickled the fairy.” They generalize from hearing one example to another. After enough examples of this kind, they might even be able to guess who did what to whom in the sentence “The gazzer mibbed the toma,” even though some of the words are literally nonsensical. The grammar must be something they discern beyond the words themselves, given that the sentences share little in common at the word level.

The meaning in language emerges through an interaction between the potential meaning of the words themselves (such as the things that the word “ate” can mean) and the meaning of the grammatical construction into which they are plugged. For example, even though “sneeze” is in the dictionary as an intransitive verb that only goes with a single actor (the one who sneezes), if one forces it into a ditransitive

construction — one able to take both a direct and indirect object — the result might be “She sneezed him the napkin,” in which “sneeze” is construed as an action of transfer (that is to say, she made the napkin go to him). The sentence shows that grammatical structure can make as strong a contribution to the meaning of the utterance as do the words. Contrast this idea with that of Chomsky, who argued there are levels of grammar that are free of meaning entirely.

The concept of the Swiss Army knife also explains language learning without any need to invoke two phenomena required by the universal grammar theory. One is a series of algebraic rules for combining symbols — a so-called core grammar hardwired in the brain. The second is a lexicon — a list of exceptions that cover all of the other idioms and idiosyncrasies of natural languages that must be learned. The problem with this dual-route approach is that some grammatical constructions are partially rule-based and also partially not — for example, “Him a presidential candidate?!” in which the subject “him” retains the form of a direct object but with the elements of the sentence not in the proper order. A native English speaker can generate an infinite variety of sentences using the same approach: “Her go to ballet?!” or “That guy a doctor?!” So the question becomes, are these utterances part of the core grammar or the list of exceptions? If they are not part of a core grammar, then they must be learned individually as separate items. But if children can learn these part-rule, part-exception utterances, then why can they not learn the rest of language the same way? In other words, why do they need universal grammar at all?

In fact, the idea of universal grammar contradicts evidence showing that children learn language through social interaction and gain practice using sentence constructions that have been created by linguistic communities over time. In some cases, we have good data on exactly how such learning happens. For example, relative clauses are quite common in the world’s languages and often derive from a meshing of separate sentences. Thus, we might say, “My brother.... He lives over in Arkansas.... He likes to play piano.” Because of various cognitive-processing mechanisms — with names such as schematization, habituation, decontextualization and automatization — these phrases evolve over long periods into a more complex construction: “My brother, who lives over in Arkansas, likes to play the piano.” Or they might turn sentences such as “I pulled the door, and it shut” gradually into “I pulled the door shut.”

What is more, we seem to have a species-specific ability to decode others’ communicative intentions — what a speaker intends to say. For example, I could say, “She gave/bequeathed/sent/loaned/sold the library some books” but not “She donated the library some books.” Recent research has shown that there are several mechanisms that lead children to constrain these types of inappropriate analogies. For example, children do not make analogies that make no sense. So they would never be tempted to say “She ate the library some books.” In addition, if children hear quite often “She donated some books to the library,” then this usage preempts the temptation to say “She donated the library some books.”

Such constraining mechanisms vastly cut down the possible analogies a child could make to those that align the communicative intentions of the person he or she is trying to understand. We all use this kind of intention reading when we understand “Can you open the door for me?” as a request for help rather than an inquiry into door-opening abilities.

Chomsky allowed for this kind of “pragmatics” — how we use language in context — in his general theory of how language worked. Given how ambiguous language is, he had to. But he appeared to treat the role of pragmatics as peripheral to the main job of grammar. In a way, the contributions from usage-based approaches have shifted the debate in the other direction to how much pragmatics can do for language before speakers need to turn to the rules of syntax.

Usage-based theories are far from offering a complete account of how language works. Meaningful generalizations that children make from hearing spoken sentences and phrases are not the whole story of how children construct sentences either — there are generalizations that make sense but are not grammatical (for example, “He disappeared the rabbit”). Out of all the possible meaningful yet ungrammatical generalizations children could make, they appear to make very few. The reason seems to be they are sensitive to the fact that the language community to which they belong conforms to a norm and communicates an idea in just “this way.” They strike a delicate balance, though, as the language of children is both creative (“I goed to the shops”) and conformative to grammatical norms (“I went to the shops”). There is much work to be done by usage-based theorists to explain how these forces interact in childhood in a way that exactly explains the path of language development.

## A Look Ahead

At the time the Chomskyan paradigm was proposed, it was a radical break from the more informal approaches prevalent at the time, and it drew attention to all the cognitive complexities involved in becoming competent at speaking and understanding language. But at the same time that theories such as Chomsky's allowed us to see new things, they also blinded us to other aspects of language. In linguistics and allied fields, many researchers are becoming ever more dissatisfied with a totally formal language approach such as universal grammar – not to mention the empirical inadequacies of the theory. Moreover, many modern researchers are also unhappy with armchair theoretical analyses, when there are large corpora of linguistic data – many now available online – that can be analyzed to test a theory.

The paradigm shift is certainly not complete, but to many it seems that a breath of fresh air has entered the field of linguistics. There are exciting new discoveries to be made by investigating the details of the world's different languages, how they are similar to and different from one another, how they change historically, and how young children acquire competence in one or more of them.

Universal grammar appears to have reached a final impasse. In its place, research on usage-based linguistics can provide a path forward for empirical studies of learning, use and historical development of the world's 6,000 languages.

## More to Explore

**Constructing a Language: A Usage-Based Theory of Language Acquisition.** Michael Tomasello. Harvard University Press, 2003.

**Constructions at Work: The Nature of Generalization in Language.** Adele Goldberg. Oxford University Press, 2006.

**Language, Usage and Cognition.** Joan Bybee. Cambridge University Press, 2010.

## From our Archives

[The “It” Factor.](#) Gary Stix; September 2014.

## ABOUT THE AUTHOR(S)



### Paul Ibbotson

Paul Ibbotson is a lecturer in language development at the Open University, based in England.



### Michael Tomasello

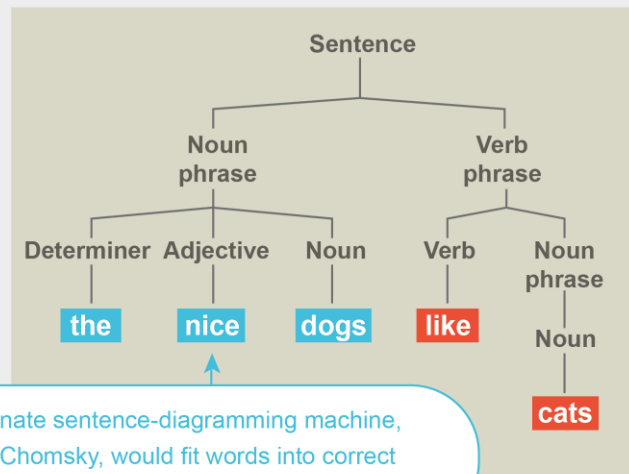
Michael Tomasello is co-director of the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, and author, most recently, of *A Natural History of Human Morality* (Harvard University Press, 2016).

## “Noam-enclature” and the New Linguistics

Noam Chomsky took the linguistics community by storm more than 50 years ago. The idea was simple. Underlying language is a set of rules innate to every child that generates grammatical sentences from the earliest age. Chomsky set out to define those rules and how they work. Without this universal grammar, he thought, it would be impossible for a child to learn any language. In the ensuing years, Chomsky's theory has gradually been challenged by new theories asserting that language is acquired as children discern patterns in the language they hear around them.

### Chomsky's Universal Grammar

Chomsky's universal grammar equipped the child with rules that worked on phrases (“the nice dogs”) and rules for transforming those phrases (“Cats are liked by the nice dogs”). The theory has evolved in recent years but still retains the essential idea that children are born with the ability to make words conform to a grammatical template.



The brain's innate sentence-diagramming machine, according to Chomsky, would fit words into correct grammatical slots—“nice” (adjective), “dogs” (noun).

### Usage-based Learning

New approaches to linguistics and psychology suggest that children's natural ability to intuit what others think, combined with powerful learning mechanisms in the developing brain, diminishes the need for a universal grammar. Through listening, the child learns patterns of usage that can be applied to different sentences. The word “food” might replace the word “ball” after the phrase “The dog wants.” Studies show that this theory of building up knowledge of word meaning and grammar approximates the way that two- and three-year-olds actually learn language.



Illustration by Lucy Reading-Ikkanda