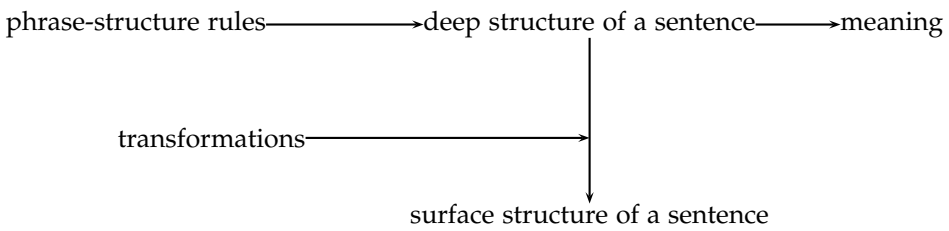
Universal Grammar: the history of an idea

John Goldsmith

March 2, 2010

The term *universal grammar* first arose (to my knowledge) in 17th century France, in the context of the Port Royal grammarians. But in trying to look back at earlier work, we have to be careful and to bear in mind that until recently, there was no effort made to separate (or even distinguish between) those characteristics of language that followed from the nature of thought and/or logic, and those characteristics that do not.

A warning about the term *deep structure*: a term that goes back to Hockett (1958), who wrote of *deep grammar*. The term was used through *Aspects of the Theory of Syntax* (Chomsky 1965), and widely misunderstood by people who did not bother to read the book. *Deep structure*, in Chomsky's model, has nothing to do with *meaning* or with what is *common across languages*. Deep structure corresponds to a kind of structurally regular description of a particular language; it is a simplification of surface patterns, and the difference between the observed (surface) patterns and this deep structure are the result of the effect of transformations, which are different in character from the phrase-structure rules which generate deep structures (in this model).



Katz-Postal-generative semantics hypothesis: deep structures are also *meaning representations*. Chomsky was unimpressed.

1. UG innatist Hypothesis: Human beings are endowed from birth (if not before!) with some forms of knowledge that did not come from or through the senses that enables them to learn human languages; without which they would not be able to learn a human language; and which is essentially the same for all human beings.

2. UG autonomy hypothesis (*domain-specificity*): The innate UG described in the first hypothesis is different in important ways from any other kinds of innate knowledge that are important in other domains of human activity and intelligence.

Be clear on the relationship between these two hypotheses! If there is a thing such as universal grammar, then it needs to be associated with the ability to infer a grammar, in the presence of linguistic input from a particular language. UG is a description of a state of readiness to learn from linguistic data: it is a theory of a particular kind of learning.

Nativism and innatism: in some quarters *nativism* is associated with a scientific reliance on genetics and/or cognitive psychology, while *innatism* goes back to Plato.

1 *Port Royal grammarians*

The Port Royal grammarians view was expressed most clearly in Grammaire générale et raisonnée de Port-Royal by Claude Lancelot and Arnauld (16xx), and Chomsky attracted a lot of attention to this forgotten work in the mid-

1960s in his *Language and Mind* and *Cartesian Linguistics*. Like most pre-modern grammarians, this school sought to explain grammar on the basis of the natural movements of the mind, and to a modern ear, their accounts beg the question in case after case. Consider chapter 24:

> La construction des mots se distingue généralement en celle de convenance, quands les mots doivent convenir ensemble, et en celle de régime, quand l'un des deux cause une variation dans l'autre. La première, pour la plus grande partie, est la même dans toutes les langues, parce que c'est une suite naturelle de ce qui est en usage presque partout pour mieux distinguer le discours....La distinction du féminin et du masculin a obligé de même de mettre en même genre le substantif et l'adjectif, ou l'un et l'autre quelque fois au neutre, dans les langues qui en ont; car ce n'est que pour cela qu'on a inventé les genres. Les verbs, de même, doivent avoir la convenance des nombres et des personnes avec les noms et les pronoms. Que s'il se rencontre quelque chose de contraire en apparence à ces règles, c'est par figure, c'est à dire, en sous-entendant quel mot, ou en considérant les pensées plutot que les mots même, comme nous le dirons ci-après.
>
> La syntaxe de régime, au contraire, est presque toute arbitraire, et par cette raison se trouve tres différente dans toutes les las langues: car les unes font les régimes pas les cas; les autres, au lieu de cas ne se servent que de petites particules qui en tiennent lieu, et qui ne marquent même que peu de ces cas...

2 *George Marsh 1860*

Marsh was a Congressman, and later a diplomat, and after the lectures (see below) he went on to become the American ambassador to Italy. He was a noted linguist of his day.

In 1858, George Perkins Marsh was invited to give a series of lectures at Columbia University in New York that would constitute "post-graduate education" on the subject of the English language, and he presented these the following autumn. Marsh was a lawyer and diplomat; he spent six years during the 1840s as a Congressman from Vermont, and shortly after presenting these lectures at Columbia, he was appointed by Abraham Lincoln to serve as the first American minister to the kingdom of Italy, a post which he retained till his death in 1882. But if he was a lawyer and diplomat, he was also an early environmentalist—and linguist: he published the first grammar in English of Icelandic in 1838.

Marsh began with the observation that Americna education had till that point largely taken mastery of English to be a subject covered early on in the educational journey —in the nursery, he said—and yet here he came to speak of language in a post-graduate course. Why should we study English now?

One answer, he noted, was that it was the decay of the language which often served as the motivation for reflection on its state. Is the highest form of the English language, he asked rhetorically, decayed, extinct, and fossilized?

What, Marsh asked, is *language*? Is it something which is unique to humans, or do we share it (perhaps in a lesser degree) with other animals? That we do not perceive a message being passed from one horse to another is not proof that they are not communicating; when we listen to a human language we do not know, we are often quite oblivious to the diversity of its sounds and contrasts:

> All will agree in denying to the lower animals the possession of language as a means of intellectual discourse ; but even this conclusion must rest upon stronger grounds than the testimony of the ear. Sounds, which to our obtuse organs appear identical, may be infinitely diversified to the acuter senses of these inferior creatures, and there is abundant evidence that they do in many instances communicate with each other by means, and in a degree, wholly inappreciable by us.

> When a whale is struck, the whole shoal, though widely dispersed, are instantly made aware of the presence of an enemy; and when the gravedigger beetle finds the carcass of a mole, he hastens to communicate the discovery to his fellows, and soon returns with his four confederates. (Conscience, *Boek der Natuur* vi.) The distinction we habitually make between articulate and inarticulate sounds, though sufficiently warranted as applied to human utterance, may be unfounded with reference to voices addressed to organizations less gross; and a wider acquaintanco with human languages often teaches us that what to the ear is, at first, a confused and inexpressive muttering, becomes, by some familiarity, an intelligible succession of significant sounds. p. 32

We must use our intelligence, then, not just our ears to decide if language is specifically human, whether it is a gift of God, so to speak, or something that humans have put together. It is, he tells us, a fundamental and inalienable part of being a human to possess a human language (p. 31ff):

> Many men pass through life without pausing to inquire whether the power of speech, of which they make hourly usage, is a faculty or an art —a gift of the Creator, or a painfully- acquired accomplishment—a natural and universal possession, or a human invention for carrying on the intercommunication essential to social life. We may answer this query, in a general way, by saying that the use of articulate language is a faculty inherent in man, though we cannot often detect any natural and necessary connection between a particular object and the vocal sound by which this or that people represents it. There can be little doubt that a colony of children, reared without hearing words uttered by those around them, would at length form for themselves a speech. What its character would be could only be determined by the method of Psammetichus,[1] an experiment too cruel to be repeated by inquirers intelligent enough to be interested in the result. It is not improbable that a language of manual sign would precede articulate words, and it may be presumed that these signs would closely resemble those so much used as a means of communication among savages, and which are, to a great extent, identical with what have been called the natural signs of the deaf-and-dumb. If you bring together two uneducated but intelligent deaf-mutes from different countries, they will at once comprehend most of each other's signs, and converse with freedom, while their respective speaking countrymen would be wholly unable to communicate at all.[2]

[1] First of three kings of the 26th Egyptian dynesty, and first documented psycholinguist (see Herodotus's *Histories*.

[2] p. 31.

> The origin of language is shrouded in the same impenetrable mystery that conceals the secrets of our primary mental and physical being. We cannot say, with some, that it is of itself an organism, but we regard it as a necessary, and, therefore, natural product of intelligent self-conscious organization.[3]

[3] p. 38.

But we will never really know about the origins of language, because everywhere we turn, the languages we saw are all of great, deep, and equal complexity:

> So far as observation goes, its structure is as complete among the most unlettered savages, and in the remotest periods, as in the golden age of Hellenic literature.[4]

[4] p. 38.

Human language is special; it is not cousin to the speech of brutes (though learning how brute animals, from birds to whales, communicate will require serious study). But is language best understood as a property of individuals? No, Marsh argues; acquisition of a language is much like an instinct.

> But though the faculty of articulate speech may be considered natural to man, it differs from most other human powers, whether organic or incorporeal, in this: that it is a faculty belonging to the race, not to the individual, and that the social condition is essential, not to its cultivation, but to its existence. Hence, its exercise is not spontaneous, or in any sense self-taught, as are all purely organic processes. Nevertheless, considered in its mode of action, the use of the mother-tongue may be regarded as an instinctive function, because it is acquired through the prompting of

natural impulses, and without any conscious, calculating effort. We retain no recollection of the process by which we learned to understand and employ our maternal speech, at least as respects that portion of it which is mastered in infant life, and not taught in the artificial form it assumes in books. In actual speaking, the movement, both physical and intellectual, is as completely automatic and unconscious as the action of the nerves, muscles, and tendons, by whose instrumentality the hand is raised or the foot thrown forward.[5]

[5] p. 39.

3 *American structuralists*

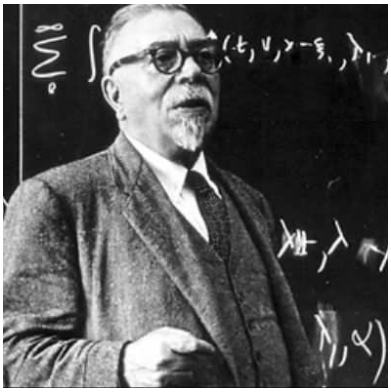
One of the defining characteristics of the American, anthropology-motivated school of linguists—the tradition of Franz Boas, Edward Sapir, and Leonard Bloomfield—was taking traditional (Indo-European-oriented classical historical) linguistics and applying the methods to analyze native American languages that had no history of analysis, nor any written form, recent or ancient. What little work existed on these languages was wildly over-influenced by the drive to give an analysis which matched the traditional categories, especially regarding parts of speech. This new *descriptive-structuralist* school rejected the impulse to ground all linguistic analysis on a comparison with Latin, Greek, and Sanskrit, and urged linguists to figure out what the appropriate categories were on the basis of the materials in the language itself.

Famous quotation by Martin Joos: Languages "can differ from each other without limit and in unpredicatble ways," and they should be approached " without any preexistent scheme of what a language must be."

4 *Post-war*

4.1 *Cybernetics*

1. Norbert Wiener. From Wiener's best-seller, *Cybernetics* 1948:



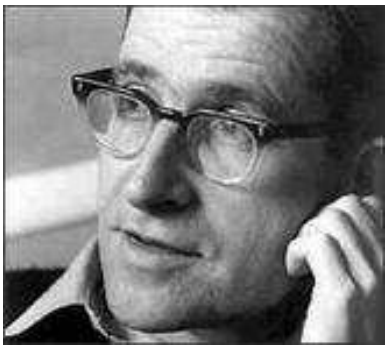
The fact that speech belongs in general to man as man, but that a particular form of speech belongs to man as a member of a particular social community, is most remarkable. . . It is quite clear that if left alone, babies will make attempts at speech. . . . It is almost equally clear that if a community of children were left out of contact with the language of their seniors through the critical speech-forming years, they would emerge with something, which crude as it might be, would be unmistakably a language.

Why is it then that chimpanzees cannot be forced to talk, and that human children cannot be forced not to? Why is it that the general tendencies to speak and the general visual and psychological aspects of language are so uniform over large groups of people, while the particular linguistic manifestation of these aspects is varied? At least partial understanding of these matters is essential to any comprehension of the language-based community. We merely state the fundamental facts by saying that in man, unlike the apes, the impulse to use some sort of language is overwhelming; but that the particular language used is a matter which has to be learned in each special case. It is apparently built into the brain itself, that we are to have a pre-occupation with codes and with the sounds of speechâ¦ The gift of speech . . . is strictly a psychological impulse, and is not the gift of speech, but the gift of the power of speech.

In other words, the block preventing young chimpanzees from learning to talk is a block which concerns the semantic and not the phonetic stage of language. The chimpanzee has simply no built-in mechanism which leads it to translate the sounds that it hears into the basis around which to unite its own ideas or into a complex mode of behavior.(p. 84)

2. Singh (1962): *Great ideas in Information Theory, Language, and Cybernetics*

p. 2-3 This inborn capacity to learn any language rather than the genetically transmitted, rigidly predetermined language patterns of the lower animals is the real mark of his humanness, because it enables him to use language in a way no other animal can. By eliminating the need for any relation or resemblance whatever between the sound (word) symbol and its meaning... man finds a limitless ocean of communication possibilities opened up to him which are denied other animals handicapped by their severely limited laryngeal-cum-cerebral prowess....One great advantage of such a free creation of systematic sound patterns charged with meanings is that by mere talk man is able to transmit almost any kind of experience vicariously, thereby creating a basis for the cumulative development of civilization and culture. Language therefore is man's open-sesame to all the treasures of the earth he has since created or found for himself.



5 *Noam Chomsky*

Chomsky has presented two very different over-all conceptions of linguistic theory, and the kind of knowledge it imputes to the language learner. The first (the *classical generative theory*) involves the notion of evaluation metric, and was at the heart of Chomsky's generative grammar from the mid 1950s (i.e., its inception) to 1979. The second has been called the *principles and parameters approach*, and has been widely applied since the early 1980s. Very little explicit discussion accompanied this considerable change.

The classical generative theory said that the goal of a linguistic theory is to specify a notion of possible grammar, and an explicit *evaluation metric*. The evaluation metric would be used to choose between alternative grammatical analyses, when both were consistent with the observed data and both fell within the limits of "possible grammars."

- Classical generative grammar proposed using length of the grammar (in a certain notation) to account for which grammar was preferred.
- Principles and parameters: a language is a set of two things: a lexicon (a set of words with grammatical information associated with each item), and a set of settings of a finite set of *parameters*; each parameter has a small number of settings.

From Chomsky 1997

p. 25: A plausible conclusion today, I think, is that the computational system is fixed and invariant; in this sense, there is only one human langauge, much as an intelligent Martian biologist would have expected. Language variation seems to be limited to the lexicon, either general properties of all lexical items or properties of specific ones, variable within a narrow range among languages....Choice among the options left open in the initial state allows for a finite number of essentially different human languages, in the Osherson-Weinstein sense, that is, up to the choice of lexical items within the restricted (but unbounded) limits allowed.

5.1 *Chomsky's 3 notions of adequacy of a grammar*

- **Observational adequacy**: a grammar of a language L is *observationally adquate* iff it generates all and only the sentences of L.
- **Descriptive adequacy**: a grammar of a language L is *descriptively adequate* iff it is adquately represents the linguistic knowledge [intuitions, perhaps?] of speakers of L.
- **Explanatory adequacy**: A grammar G of a language L is *explanatorily adequate* iff it is selected by a larger theory that provides an explicit account of why G

is preferable to any and all other grammars for L. The principles of this larger theory constitute UG (Universal Grammar).

Chomsky's early, classical theory of generative grammar was that the crucial element of UG would be an evaluation metric, which would assign a length to a grammar, and the grammar selected by the UG would be the shortest grammar consistent with the data from language L.

Principles and parameters: this "was based on the idea that the format consists of invariatn principles and a "switch-box" of parameters. . . . The switches can be set to one or another value on the basis of fairly elementary experience. A choice of paramters settings determines a language. The approach largely emerged from intensive study of a range of languages, but as in the early days of generative grammar, it was also suggested by developments in biology—in this case, François Jacob's ideas about how slight changes in the timing and hierarchy of regulatory mechanisms might yield grat superficial differences." (Chomsky 2007, p. 17)

6 *Late 20th century views*

General observations: Chomsky: "Diversity of language provides an upper bound on what may be attributed UG: it cannot be so restricted as to exclude attested languages. Poverty of stimulus (POS) considerations provide a lower bound: UG must be at least rich enough to account for the fact that internal languages are attained."

6.1 *Arguments for innatism*

See Chomsky, *Of Minds and Language* 2007—assigned reading.

One group of these arguments have been called *poverty of the stimulus* arguments: the speaker shows firm certainty for certain facts of his language which are not directly supported from evidence he had access to while he was learning the language. The data is too impoverished to account for grammatical conclusions. This argument is especially compelling if the true-but-underetermined conclusions are consistent across all languages. If they are not, then the conclusion may be that we (simply) do *not know yet* how the learner drew the conclusion.

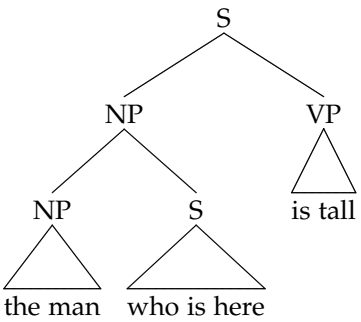
I think that a slightly different view yields more insight into why Chomsky and those who follow him in this hold to this firm, substative innatist view. In order to provide a formal (and not-meaning-based) account of syntactic properties, it has proven necessary to state relatively complex conditions. If linguistics is taken as a model of the human mind, then, these conditions and the formalisms that allows their specification must be part of that mind: but it is difficult to imagine that this could be learned.

1. The data from which the child learns language presents no clear evidence that it must be interpreted hierarchically, but the learner always assumes that it is hierarchical anyway. That assumption is prior or innate knowledge of the child. Chomsky's example: why isn't the rule: *Move the first auxiliary verb to the front of the sentence*?

(a) {The man [who is here]} is tall.

(b) Is {the man [who is here]} [e] tall?

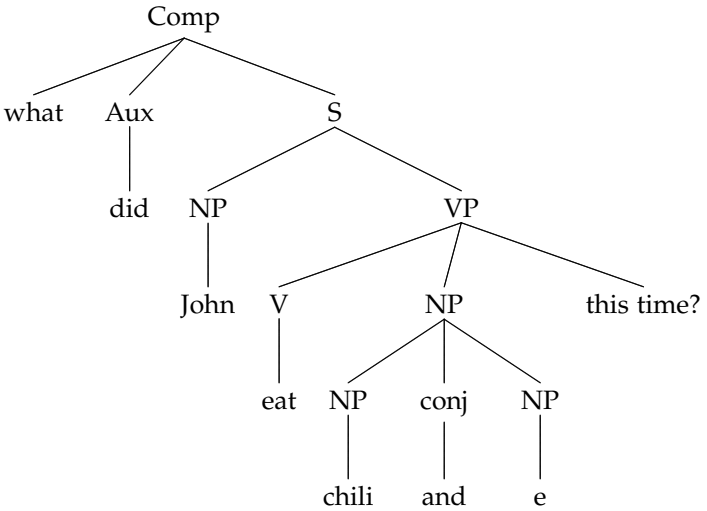
(c) *Is {the man [who [e] here]} is tall?



2. All children learn language (unlike mathematics or music)

3. Children pass through similar stages in their acquisitions of the most varied of languages.

4. Specific proposed universal constraints, such as Haj Ross's Coordinate Structure constraint: *in a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct*. (1967). Suppose John gets quite sick when he eats chili along with various foods, such as cornbread, chicken, and carrots. If we heard that John had gotten sick, wouldn't it be natural to want to ask, *What did John eat chili and [e] this time?* But that is ungrammatical, and not just in English, but in language and language.



6.2 *Arguments against innatism: Evans and Levinson*

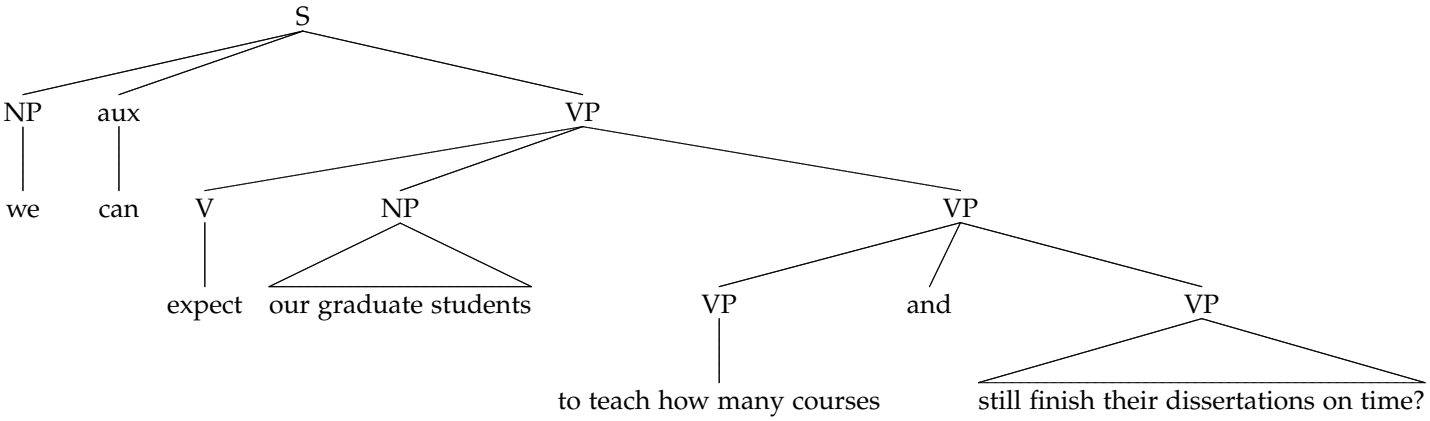
See Evans and Levinson, *Behavioral and Brain Sciences* 2009—assigned reading.

They begin with their main point:

> Languages are much more diverse in structure than cognitive scientists generally appreciate... The claims of Unviersal Grammar, we argue here, are either empirically false, unfalsifiable, or misleading in that they refer to tendencies rather than strict universals. Structural differences should instead be accepted for what they are, and integrated into a new approach to language and cognition that places diversity at cetre stage." (429)
>
> How did this widespread misconception of language uniformity come about? [Perhaps ethnocentrism...]Unfortunate sociological splits in the field have left generative and typological linguists with completely different views of what is proven science, without shared rules of argumentation that would allow them to resolve the issue—and in dialogue with cognitive scientists it has been the generativists who have been taken as representing the dominant view. As a result, Chomsky's notion of UG has been mistaken, not for what it is—namely, the programmatic label for whatever it turns out to be that all children bring to learning a language—but for a set of substantial research findings about what all languages have in common.

1. The argument from *structure-dependence* assumes that transformations apply to strings of words (with their bracketed or hierarchical structure). If instead transformations apply to phrase-structure rules to create alternative phrase-structure rules, then the argument does not go through. $S \rightarrow AuxNPVP$. Can such an analysis be articulated and defended?

2. No suggested piece of innate knowledge has advanced from *proposed* to *established* (even if we set relatively casual standards for *established*). For example, the Coordinate Structure Constraint: Ross had already observed that the constraint did not hold if we extract in an *across-the-board* fashion: *Who was it that John offered to invite [e] and then failed to contact [e]?* Or if the first conjunct is auxiliary-like: *What do you suppose Henry went and did [e] next?* Later others noted that some extractions from the first conjunct were fine too: *How many courses can we expect our graduate students to teach and still finish their dissertations on time?*



3. Janet Pierrehumbert (2000):

> Languages can differ systematically in arbitrarily fine phonetic detail. This means we do not want to think about universal phonetic categories, but rather about universal phonetic resources, which are organized and harnessed by the cognitive system... The vowel space—a continuous physical space rendered useful by the connection it establishes between articulation and perception—is also a physical resource. Cultures differ in the way they divide up and use this physical resource.

4. We cannot provide a scientific account for a human capacity (the language faculty); this inability on our part does not constitute an argument that no scientific account is possible and therefore the observed knowledge must be innate. That constitutes inappropriate scientific defeatism, not scientific discovery.