

Noam Chomsky's

UNIVERSAL GRAMMAR

&

The Current State of The Theory

Samuel Pearce
2021

Abstract

Noam Chomsky's Theory of Universal Grammar has been the defining theory behind linguistics for the past several decades. Many linguists have based their career around how they stand in relation to it, and no matter one's stance on the matter, it is impossible to refute the Theory's monumental impact on the field as a whole. This paper aims to outline — very briefly — what the theory entails, how it has evolved since it was first proposed in the 1960s, and what the most current consensus on the matter is.

Contents

1	Universal Grammar	3
2	Generative Grammar	3
3	Government/Binding Model	4
4	X-Bar Theory	5
5	Principles & Parameters Theory	6
6	Minimalist Program	7
7	Criticism	7
8	My Impression	7

1 Universal Grammar

To begin with, what is universal grammar exactly, and what does it entail? Well, contrary to what I originally thought, it is not a fully-developed linguistic grammar that applies to all human languages. The theory of universal grammar mainly consists of the hypothesis that language is an innate human faculty that is defined — at least, to some extent — by our biology and genetics, and a developing structure of grammar that outlines the most basic system that applies to all natural languages. One of the well-known arguments Chomsky presents in favour of this hypothesis is that young children can understand grammatical concepts about the language they speak without having been taught them directly. (Cook and Newson, 2007, p. 26)

At its base, UG-theory suggests there is some kind of “computational system” that converts the external signals (sounds, symbols, sign-language, etc.) into internal ideas. (Cook and Newson, 2007, p. 5) Most of Chomsky’s work in UG was creating and developing a formal system to analyse syntax of varying languages to compare which elements might be universal. The ideas and structures we use to analyse human communication and form ideas about how such an internal system might be structured has been discussed and evolved over the past nearly half of a century:

“The UG Theory claims to be a scientific theory based on solid evidence about language. As such, it is always progressing towards better explanations for language knowledge....”

(Cook and Newson, 2007, p. 27)

The theory split into two main branches, each focusing on one end of the computational system: “E-Language” and “I-Language”. E-Language is primarily concerned with the real, physical manifestation of language and gathers large samples to analyse. The I-Linguists focus on mapping out how these ideas are stored in the brain, with the details of how they are expressed mostly irrelevant. To summarise: E-Language “is concerned with what people have done”, while I-Language “is concerned with what they could do” (Cook and Newson, 2007, p. 14)

One of the first things one would think of analysing when proposing a theory of *universal* grammar are so-called linguistic universals: features that have the same structure throughout all human languages. Greenbergian universals, linguistic features such as syntax structure or movement rules that appear in all natural languages (Greenberg, 1969), differ from Chomskyan universals in that Chomskyan universals needn’t manifest in every language; “No language violates a universal principle (the language simply may not use the principle in a particular context)” (Cook and Newson, 2007, p. 23)

2 Generative Grammar

The name of generative grammar can be misleading and is often misunderstood, at its core, it’s a description of a language that uses a very explicit syntax which accounts for every possible construction. Chomsky himself defined it as such:

“When we speak of the linguist’s grammar as a ‘generative grammar’ we mean only that it is sufficiently explicit to determine how sentences of the language are in fact characterised by the grammar”

Chomsky (1980, p. 220)

It is based on the idea of building up sentence syntax the same way one can with a mathematical grammar or a programming language; possible syntax trees are described with recursive rewrite rules such as:

$$S \rightarrow NP VP$$

$$VP \rightarrow V NP$$

$$NP \rightarrow Det N$$

In the above example a sentence (S) is defined as a noun-phrase (NP) plus a verb-phrase (VP), a verb-phrase consists of a verb (V) and a noun-phrase, and a noun-phrase is a determiner (Det) plus a noun (N). With these rules you can construct some of the many possible grammatically valid sentences for English. (Cook and Newson, 2007, p. 32)

Though this was just the beginning of a new way of analysing the syntax structures of sentences which lead to many further developments.

3 Government/Binding Model

In the new government/binding model, the simple building blocks of language which were still being defined were stuck together modularly to create a model of how the computational system could possibly be structured. At its base, the structure has two layers: the D-structure (deep structure) and the S-structure (surface structure). The D-structure reflects the base grammatical structure of the sentence, influenced by the lexicon and phrase structure rules explained above, while the S-structure is what is actually said after being adjusted by movement rules. (Cook and Newson, 2007, p. 61) For example, in the following sentence, the “whom” would be the object of the sentence — coming after the verb, as usual — in the D-Structure, but because it’s a question, the object is moved to the front of the sentence:

* You did see whom. \rightarrow Whom did you see?

This was the basis of the GB theory which would continue to be developed with new additions to help handle all cases seen in human languages that didn’t already meet the basic model. One of the first iterations of the model, as described above, is illustrated in the following figure from Cook and Newson (2007, p. 62)

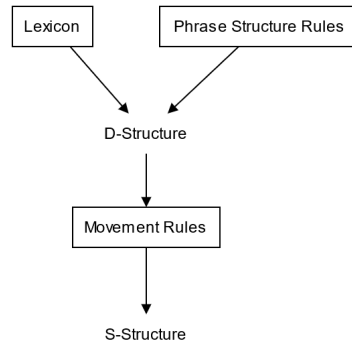


Fig. 1: One of the simplest forms of the GB-Model

4 X-Bar Theory

In the earlier example we stated that a verb-phrase was defined as a verb and a noun-phrase, but some verbs are intransitive and don't take an object. While we could simply define two possible verb-phrase trees — one with a object and one without — this information is based on the word itself and this causes an unnecessary redundancy. Chomsky's rewrite rules were then expanded upon with the introduction of X-bar notation to remove this redundancy; a phrase could be defined using only the following rewrite rules:

$$X' \rightarrow X (YP)^1$$

$$X'' \rightarrow (YP)^2 X'$$

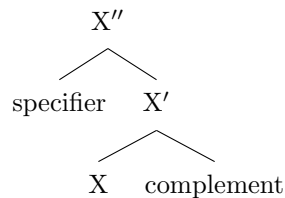


Fig. 2: The X-Bar rules displayed as a tree

Substitute X and Y with any of the four lexical categories in the theory — Noun (N), Verb (V), Adjective (A), Preposition (P) — and you have the structure of that phrase. This system can represent any phrase possible by recursively adding more elements to the tree when necessary. (Cook and Newson, 2007, p. 66)

¹The complement is optional; it's determined by the head

²The specifier is also optional, but isn't selected by the head

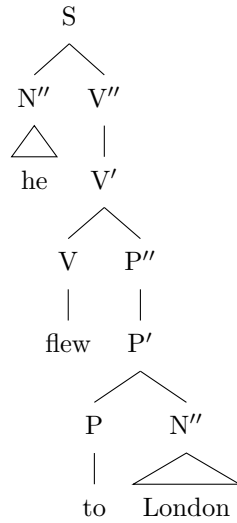


Fig. 3: An example of recursive phrases from Cook and Newson (2007, p. 68)

5 Principles & Parameters Theory

Principles and Parameters (P&P Theory) is a continuation of the GB-model which more accurately represents the more general nature of GB theory.

“These modules of language stand alongside many others ... Determination of the nature of these and other systems is a common project, not specific to this particular conception of the nature of language and its use”

Chomsky (2014)

Principles replaced the idea of fixed grammatical rules with a more generic set of constraints that apply to all constructions in a language as opposed to very specific ones, as with ordinary rules. (Cook and Newson, 2007, p. 35) An example of one such principle is the locality principle:

“**Locality:** this principle is a property of linguistic processes which restricts their application to a limited part of the sentence. This then forces movements in all languages to be local: *they must be short*.”

Cook and Newson (2007, p. 41)

Parameters are what account for the wide variation we see between languages; if the only thing governing languages were a set list of principles, they would be far more similar than we observe them to be. Cook and Newson (2007, p. 44) use the head parameter to illustrate how parameters differ from principles. The head — the most important word in a phrase, which selects what kind of complements may be used with it — may come before or after its complement (or either side such as with French adjectives), depending on the language it’s used in; While it is a principle that a phrase must *have* a head,

where it goes is decided by the head parameter and differs from one language to another.

With P&P theory and X-bar notation, the GB-model was expanded to include more modules over the course of many years of study introducing new features that needed to be accounted for, leading to the following version of the model:

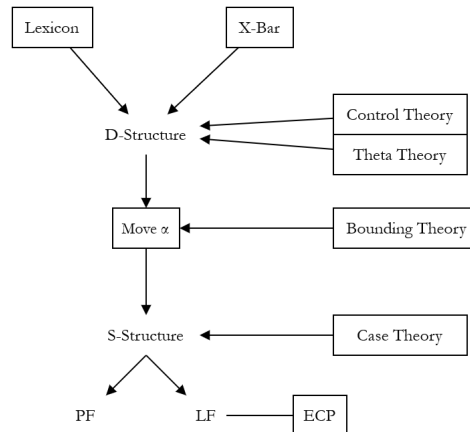


Fig. 4: The latest structure of the GB-model (Cook and Newson, 2007, p. 181)

6 Minimalist Program

The minimalist program is the name for the latest (since 2000) development out of the GB and P&P Theories. it espouses any claims at determinacy by its name; it's a program at the moment, not a developed theory; it's the approach being taken to explore the most recent and least-well understood linguistic concepts. (Cook and Newson, 2007, p. 242) So far, the minimalist program has had very little impact on language acquisition research specifically: "Most L1 researchers seem quite happy to continue using a version of the P&P model that involves little of the MP" (Cook and Newson, 2007, p. 219)

The minimalist program had three main phases of its development: in the beginning, it focused on finding the most foundational principles of the worlds languages and the principles which made up for the discrepancies between them. After that, the linguists began examining how well the phonetic form (PF) and the computational system interfaced. a large portion of the GB-model was demolished for MP in order to reconstruct the syntax in a manner which tailored it to the most minimal linguistic ideas. (Cook and Newson, 2007, p. 3) Since then, the phases model has been the latest production of the minimalist program. One can read about the phases model from *The Architecture of Language*, Carr (2005)

7 Criticism

Chomsky's theory of Universal Grammar has been the subject of relentless criticism since it first came about in the 1960s. One of the most common arguments

against it, which targets the very foundation of the concept as a whole, criticises the assumption that psychological theories can be evaluated in the same way physical ones can; some, such as Quine (1960), claimed they were fundamentally different, and therefore, could not be evaluated equally. Chomsky rebutted this by laying out his view that there is no current proof that one cannot distinguish between two psychological theories as well as with physical ones. (Cipriani, p. 13)

Another point often brought up is that the principles in P&P theory must limit the scope of human language, no matter how extensive the parameters alter their representations. Though, this can be fairly easily dismissed as it is an inherent property of the theory; if you wish to make a list of principles common to all languages, it must have a fixed boundary in order to have any meaningful value. (Cipriani, p. 13)

A very strong criticism levied against UG is that it has no limit to how many exceptions it can have while still being considered correct. In other fields such as mathematics, you would have to abandon a theory once it exceeded the ideal limit of exceptions, but UG, according to its own definition, is always correct. The consequence of this is that UG theory is not falsifiable, and therefore — according to some — unscientific. (Cipriani, p. 14)

8 My Impression

Personally, I find the last of the example criticisms very compelling; as I read through Cook and Newson (2007), I often found myself lamenting the rapid loss of elegance in any new concept introduced. At the time, I couldn't put into words the persistent feeling I had that this *couldn't* be universal; there were far too many exceptions and specific rules and it seemed to only apply to a handful of the most well-known languages like English, German, Japanese, etc. Likely this is partially due to the very brief introduction I have been given, but I still feel it's unproductive to analyse something that, we may never be able to prove. I'm very appreciative of the many positive impacts UG's study has had on the entire field of linguistics and computer-science, but I can't help feel it's unscientific to make claims about the extremely subjective qualia which permeate every discussion of human language. It makes a fantastic *model* of human languages in a very elegant manner, but — at least for now — it *can* only ever be a model.

Another concept troubling me while studying this is that it seems to miss a question that seems fundamental to the topic: are natural human languages spoken around the world really representative of the full extent of human language faculty (assuming it is a biological faculty)? How can we tell where the border of this in-built faculty and simple logical deduction is? It's possible that we could learn more complex languages, or simply completely different languages from what seems possible given all current knowledge of human language, while this is taken up by other parts of the brain and it would be indistinguishable from our regular language faculty. Essentially, Does this simply also count as part of UG, or is this the brain simply adapting to a new language by using it's other logical faculties?

These questions were prompted by my interest in constructed languages (conlangs) such as Toki Pona (Lang, 2014) and Ithkuil (Quijada, 2011) which

test the extreme limits of our ability to parse language. I would be interested to know how these fit into UG theory, if they do at all, because their ability to test what we can interpret as “language” seems to suggest to me that while we are probably equipped to develop, speak and understand language by evolution, it might be far more open-ended than previously thought. Personally, I think the negative stigma still attached to conlangs and the prevailing opinion that they aren’t “real languages”, and therefore not worth studying, might be severely hindering linguistic progress. In my opinion, until we can fully accept conlangs as the powerful avenues of linguistic exploration they really are, making any significant progress towards understanding language and human nature as a whole will be stunted.

References

- Philip Carr. Noam chomsky, the architecture of language. *Journal of Linguistics*, 41(2):445—452, 2005. doi: 10.1017/S002222670523339X.
- Noam Chomsky. Rules and representations. *Behavioral and Brain Sciences*, 3(1), 1980. doi: 10.1017/S0140525X00001515.
- Noam Chomsky. *The Minimalist Program*. 01 2014. ISBN 9780262327282. doi: 10.7551/mitpress/9780262527347.001.0001.
- Enrico Cipriani. The generative grammar between philosophy and science. *European Journal of Literature and Linguistics*, pages 12–16.
- V. J. Cook and Mark Newson. *Chomsky’s Universal Grammar: An Introduction*. Blackwell Publishing, 3 edition, 2007.
- Joseph H. Greenberg. Language universals: A research frontier. *Science*, 166(3904):473–478, 1969. ISSN 0036-8075. doi: 10.1126/science.166.3904.473. URL <https://science.sciencemag.org/content/166/3904/473>.
- S. Lang. *Toki Pona: The Language of Good*. Great Britain, 2014. ISBN 9780978292300. URL <https://books.google.ch/books?id=5P0ZjwEACAAJ>.
- J. Quijada. *A Grammar of the Ithkuil Language*. 2011. URL <https://books.google.ch/books?id=HkNTMwEACAAJ>.
- Willard Van Orman Quine. *Word and object: An inquiry into the linguistic mechanisms of objective reference*. 1960. ISBN 0262670011.