

Prologue: Is Bigger Better?

ANNIE: *Anything you can do, I can do better.*

I can do anything better than you.

FRANK: *No you can't.*

ANNIE: *Yes, I can.*

FRANK: *No, you can't.*

ANNIE: *Yes, I can.*

FRANK: *No, you can't.*

ANNIE: *Yes, I can, yes, I can!*

—"Anything You Can Do," Irving Berlin, *Annie Get Your Gun*, 1950

The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin.

—Rich Sutton, "The Bitter Lesson," 2019

A Nobel Interview in the Wee Hours, October 8, 2024

On October 8, 2024, at 11:45 a.m. Central European Time, the Physics Nobel Prize Committee announced that this year's prize was to be awarded to Geoffrey Hinton—a professor of computer science at the University of Toronto and laureate of many prestigious prizes—alongside Princeton's John Hopfield. The Committee noted that Hinton had "used the Hopfield network as the foundation for a new network that uses a different method: the *Boltzmann machine*. This can learn to recognise characteristic elements in a given type of data. Hinton used tools from statistical physics, the science of systems built from many similar components. The machine is trained by feeding it examples that are very likely to arise when the machine is run. The Boltzmann machine can be used to classify images or

create new examples of the type of pattern on which it was trained. Hinton has built upon this work, helping initiate the current explosive development of machine learning.”¹

Shortly after the announcement, Hinton, who was housed in a “cheap hotel in California,” as he described it, received another call, this time from Adam Smith of the Nobel Prize website. It was about 3 a.m. Pacific Time, and Smith was calling for the customary “First Reactions” brief interview.² The interview went quite well. Here is its final minute:

SMITH: *So, for instance with the large language models, the thing that I suppose contributes to your fear is you feel that these models are much closer to understanding than a lot of people say. When it comes to the impact of the Nobel Prize in this area, do you think it will make a difference?*

HINTON: *Yes, I think it will make a difference. Hopefully, it'll make me more credible when I say these things really do understand what they're saying.*

SMITH: *Do you worry that people don't take you seriously?*

HINTON: *So, there is a whole school of linguistics that comes from Chomsky that thinks that it's complete nonsense to say these things understand, that they don't process language at all in the same way as we do. I think that school is wrong. I think it's clear now that neural nets are much better at processing language than anything ever produced by the Chomsky School of Linguistics. But there's still a lot of debate about that, particularly among linguists.*

Many found this text somewhat shocking. Is Hinton's field of view from the very top of the world marred by yet another peak that must be conquered? Is the superiority of large language models (LLMs) so fragile that it must be stipulated at this sweet moment? What is the nature of the debate Hinton is alluding to? Is it not overly ambitious to expect, indeed to claim, that a physical model that classifies images well can also understand language?

This book is an extended meditation on these questions. A scientific account of the linguistic abilities we humans seem to possess has long been sought after. The engineering effort to build talking machines is familiar to every reader of this book, whether through its best current incarnation, ChatGPT, or through its tributaries and offspring. Many readers may also be aware of Noam Chomsky's words, before he went silent, to the effect that like bulldozers, deep learning devices such as Google Translate are useful, but they tell you nothing about what human language is. Whence this debate, and what exactly is it about? In this book, I address these questions in detail, describe the state of the art on both sides, explain the ideas

behind each approach, highlight their accomplishments and problems, and conclude by making some proposals for a better future.

Telling a Good Story

First, we must listen to the story as told by both sides—we humans need stories that motivate us to persist and persevere. And here, there are two stories: one, *linguistic theory*, is told by scientists; the other, *LLMs*, is told by engineers. Linguistic theory and linguistics as a field have had major accomplishments over the past 70 years, but these have not lent themselves easily to the development of practical tools, and they have often been scorned as being useless; LLMs' language application developers and AI as a field build great devices and tools, but they have found themselves accused of putting the practical before the theoretical, of having biases (gender, race, and perhaps other), of condoning exploitative conduct, and of developing monstrous machines that are a threat to the future of humanity.³

Each story claims exclusive rights, as we shall see. Which, if either, should we believe? This book explores the current state of scientific and technological inquiries into human language. The public discussion of this hot topic has touched on many important points,⁴ but naturally, certain critical questions have remained veiled. Using nontechnical language, I first offer a tutorial in which I present basic facts and explain basic issues underlying each position, and I follow by putting these critical questions on the table, in an attempt to launch an open discussion. As a cognitive neuroscientist who is also trained in formal linguistics, I focus on highly specific issues and bring up—in nontechnical language—a host of novel arguments that rely on clearcut empirical considerations and experimental results. I also try to come up with some concrete answers and some modest proposals for a better future. Let me offer a brief sketch of the two stories.

The Linguistic Story

Up until the end of World War II, the academic study of language was mostly engaged with exciting stories of the past or the exotic: European philologists studied Old Church Slavonic palimpsests, the declension of nouns in

Attic Greek, and the etymology of *strozzapreti*, that famed Northern Italian pasta, while their North American counterparts investigated the dialects of Ojibwe or the concept of time in Hopi. These collections of observations and facts, drawn from varied human activities, fascinated scholars of old texts, as well as cultural historians and anthropologists.

After the war, the language sciences underwent dramatic changes. Mathematical tools gave way to grammatical formalisms, which paved the way for Noam Chomsky's dazzling appearance on the intellectual scene. Chomsky had a new story to tell, one that revived ancient questions: human language, he said, is not just a bag of words but a formal system of grammar rules of various types. We as speakers know these rules, which leads to the conclusion that they are part of our mental existence. Linguistics is therefore the study of a critical ingredient of human mental structure. In making this argument, Chomsky placed linguistics at the heart of the philosophical and psychological investigation of the human mind, which naturally connected to questions of cognitive development in childhood, the dissolution of cognitive faculties subsequent to brain disease, and a host of related psycho- and neurobiological issues.

Capturing the imagination of many, Chomsky proposed a new research program, and with the help of Morris Halle, Jerome Wiesner, and several others, a graduate program in the new linguistics was founded at MIT in 1961, with a focus on formal linguistic theory in all its manifestations, from phonetics to morphophonology and syntax and, later, to semantics and pragmatics. Under Chomsky and Halle's tutelage, the theory developed in important and interesting ways, both substantive and methodological, discovering a huge body of new facts and puzzles about language and proposing a highly structured theory with myriad abstract generalizations that explain these puzzles. Generations of linguists have brought the theory to where it is today—light years ahead of where it started. These advances have enhanced our understanding of language, our humanity, and, to an extent, the linguistic brain, shedding light on certain intricate patterns of experimental results. And yet, hardly any of the highly sophisticated formal theories that theoretical linguists have developed are successfully implemented in language technologies. The industry thus turned away from the story told by linguists and connected to another exciting story.

The Language Engineering Story

Since at least the 1940s, neurophysiologists, mathematicians, and engineers have tried to realize the dream of modeling the central nervous system with neuron-like networks that can learn; in parallel, first attempts at creating mechanical translation machines were made. Three-quarters of a century later, and following many false starts, the two were combined in LLMs, whose crown jewel, ChatGPT, shocked the world. This amazing machine, which seems to have been created mostly by smart Millennials who are good at math, has been immediately put to use by Gen Z kids, and actually by everyone else too. As every reader of this book knows, ChatGPT makes you feel that you're conversing with a smart and knowledgeable friend, a human just like you, though a faster and better informed one. ChatGPT in its various incarnations is thus the present chapter of several long-standing traditions, which were rolled into a new, attractive story—an artificial simulation of the human brain that behaves like a human does and that, so the story goes, accounts for findings from “wet” brain experiments. This Herculean R&D effort by industrial agents has thus raised hopes that an understanding of one of humanity's greatest mysteries—our ability to communicate through language—is within reach.

The Grim Way Things Are

These two approaches to language and communication have turned out to be vastly different. The stories they tell rely on different mathematical machineries, they find different language tasks important, their products are quite distinct, and if they make new discoveries, these are of very different kinds. The juxtaposition between them raises several questions: How adequate is the account of language given by each? Can linguistic theory help the development of LLMs (e.g., by trimming massive power-hungry computations, thereby saving them time and electricity—a major issue that has led some to suspect LLM development is about to hit a wall)? Can LLMs help linguistic theory (e.g., by delivering better implementations of language acquisition and processing algorithms)? Haven't we fallen victim to an *implementation bias*, which prefers running computer programs to theories that may have advantageous properties other than “runnability”?

The name of this book suggests a preference (though not a bias) on my part: as a scientist, my main inclination is to look for the scientific merits of a theoretical account. Technological victories are important and extremely hard to accomplish, but they do not necessarily explain our humanity. Throughout the book, I will try to draw a clear line between technology and science (although admittedly, such a line may not always exist). The questions to be addressed here, then, are hard and subtle, and I will try to provide clear and detailed answers. Toward the end, I will touch on further puzzles that have more of a sociological flavor: Do the two approaches currently benefit from each other's toolbox? Is there cross talk between the two communities? Can a combined, interwoven account work better? These questions are critical to both science and industry. Beyond answers, I will try to make some modest proposals. But I can already provide an unhappy hint: something seems to have gone wrong, as sadly, there is hardly any dialogue between the language engineering and theoretical linguistics communities. If anything, there is mutual scorn and bile, coming out of leaders such as Geoffrey Hinton, who as noted has won the 2024 Nobel Prize in Physics and who, in addition to many others, recently stated that "Chomsky managed to convince his followers that language wasn't learned" (a plainly false attribution) and has repeatedly called Chomsky "crazy."⁵ Hinton and others have taken stabs at Chomsky but not so much at the theory he and many subsequent generations of theoretical linguists have developed. On the "other side," the same Noam Chomsky—my revered mentor—also had his share of sweeping dismissals of AI.⁶

These dismissive and divisive comments, coming from leaders on both sides of the aisle, are most regrettable and not helpful. No wonder, then, that hardly any joint work is ever initiated, to everyone's detriment. To me, a dialogue is prerequisite to discovery and progress, hence a must. "The key word here is *encounter*," writes food historian Massimo Montanari in his recent *Short History of Spaghetti with Tomato Sauce*. "The more numerous and more interesting the encounters, the richer the result, the stronger and more robust the plant." I could not agree more.

The Critique in Brief

The quote from ChatGPT at the beginning of this book is coherent and grammatically well-formed. Indeed, LLMs are touted as "revolutionary" devices,

whose development is an unprecedented “paradigm shift.” But this quote also turned out to be fabricated, “hallucinated” by the machine. While I will not offer a general critique of AI (such as the one in Gary Marcus’s recent *Taming Silicon Valley*),⁷ I will reflect on such oddities and their relevance to the evaluation of LLMs as models for human language behavior. The hard look I will take at these models from a (neuro)linguist’s perspective will expose deeply deficient practices: the preference of prediction to structured and constrained explanation; the belief in brute force computation (eloquently expressed by Sutton’s quote above), even if it entails the construction of hyperpowerful machines whose nontransparent internal structure is difficult to analyze, and consequently might be doing much unnecessary computation that consumes unimaginable amounts of precious energy; the unreasoned exclusion of apparently relevant datasets; the existence of concealed biases; deficient evaluation abilities through inadequate benchmarks; absence of theoretical constraints; and a frequent failure to consider alternatives.

I will not shy away from criticizing linguists’ practices, which also tend at times to be problematic, albeit perhaps to a lesser degree. In some though not all theoretical linguistic communities, one notices a lack of interest in highly relevant empirical data, typically in the results of psycholinguistic and neurolinguistic experiments; moreover, as a field, theoretical linguistics has maintained weak standards regarding quantitative methods for data analysis. Finally, linguists have been slow in gearing themselves toward implementations and useful practical applications. All of these have also been quite detrimental to the development of the field.

The discussion to come, then, will cover both sides. It will be bidirectional, though not symmetric. I will conclude by sketching a cooperative horizon, which promises to advance both science and engineering.

An Alternative and Its Merits

A critique of an approach is most valuable and compelling when accompanied by an alternative proposal, or at least a sketch thereof. In the final chapters of the book, I will present an approach that is both empirically grounded and based on methodological considerations. After demonstrating what has gone wrong, I will make a general proposal: to replace the current gigantic LLMs with small, highly constrained and modular networks.

Then, I will show how this proposal for the modeling of language is very much in the spirit of current modeling of mammalian vision and in keeping with what we know about the mosaic of language areas in the human brain.

What's in the Box?

Let me now go briefly over the book's content, by parts and chapters, and sketch the issues I discuss. Following this prologue, the book has four two chapter parts and an epilogue. In Part I, "Where We Are: The Language Sciences," I tell the story of generative grammar, its origins, development, and current state. Chapter 1, "Grammars, Old and New" I describe the beginning of grammar, its development over the ages up to the 1950s, and the intellectual climate that gave rise to Noam Chomsky's earliest proposals of transformational generative grammar (TGG), which I sketch and rationalize. Chapter 2, "Hints from Present-Day Linguistic Theory," gives a quick nontechnical tour of TGG's current state of the art, at a level that covers multiple (at times competing) linguistic frameworks. In both chapters, I tell the story of the theory and its main insights via presentations of specific issues in syntax and semantics, which concerns utterances that should be close to any speaker's (and reader's) heart and mind. I try to emphasize the problems, the logical structure of their proposed solutions, and the implications these may have for a general explanation of the human linguistic ability. Importantly, I present the theory while minimizing technical issues. This approach not only has the advantage of being inclusive of multiple linguistic perspectives and styles, but also, it assumes no prior knowledge of linguistics. Part II, "Where We Are: Language Technologies," gives a similar treatment to LLMs. Chapter 3, "From Early Perceptrons to Deep Learning," describes the history of neural networks from its anatomic-physiological beginnings in the early twentieth century to the work of the first modelers of neurons in the 1940s through the late 1950s' learning rule and all the way up to deep learning and the dramatic appearance of AlexNet in 2012. Then chapter 4, "Dreams About Talking Machines: Current LLMs," explains (again, nontechnically and requiring no prior knowledge) what networks are and what they can do. Part III, "Getting Under the Hood," is a juxtaposition, beginning with chapter 5, "A Duel: Science Versus Technology,"

in which the conceptual underpinnings of the theory of TGG is compared to LLMs. In it, issues such as the Innateness Hypothesis, constrained versus unconstrained rule systems, and other conceptual issues are discussed, with concrete case studies. Chapter 6, “Poking Linguistic Holes in LLMs (ChatGPT Included),” takes a hard look at ChatGPT and its kin. Unlike many, I do not engage in a cat and mouse chase of the errors, hallucinations, confabulations, and biases of this impressive machine. Rather, I try to locate fundamental problems, and reflect on why they are there (hint: lack of constraints or linguistic structure in the underlying networks). Part IV, “The Brain’s Language Code,” is dedicated to linguistic theory and LLMs in the context of the human brain. This part is purely about science—about the best account of the way the human brain talks and understands. Its first chapter 7, “A Mosaic of Neurolinguistic Modules,” reviews what we know about the functional structure of brain parts that are entrusted with linguistic analysis. Providing a linguist’s perspective on brain–language relations, this chapter presents weighty hints—from experiments with brain-damaged individuals, from functional neuroimaging in healthy participants, and from direct measurements of brain activity—in favor of the view that the modular nature of linguistic theory is reflected directly in brain structure. That is, specific brain parts carry out specific linguistic computations, in a manner dictated by TGG. The counterpoint chapter 8, “Modeling the Linguistic Brain with LLMs,” presents the LLM perspective, by which a certain class of neural network models is said to account for all language performance, as measured directly from the human brain. Here, too, the general view is bolstered by experimental results from brain measures that are similar to the previous ones, although the conclusion is quite different. The two perspectives may appear incommensurate, and each has problems, which will be discussed.

In the epilogue, “Shall We Work Together?,” I consider ways for us to move forward, beyond division, sectarianism, and powerplay. Instead of two competing factions playing against each other, we need to team up in a joint effort to understand the nature of human linguistic ability. And while I can already disclose that my account will not be entirely symmetric and will be more inclined toward the linguistic side, it is clear to me that neural modeling and implementation are critical ingredients of a theory that strives to give us deeper understanding of brain mechanisms for language.

How to Read This Book

I have aimed to reach a broad spectrum of readers: people interested in philosophy, science, and technology, curious readers who have heard about modern linguistics and about Noam Chomsky but have never understood what the fuss was about; engineering students who have taken or plan to take a course in natural language processing; linguists who fear that their discipline may soon be erased off the academic map; neurologists and neurosurgeons who have seen the wonders of the talking brain, as well as speech and language therapists who try to fix it when it fails; engineers and developers who may have followed some of the debates on X (formerly Twitter), YouTube, and other media, and want to build better language bots; and, of course, my academic colleagues from all relevant disciplines. There's something here for everyone, I hope.

The book touches on technical issues and tries to do so in a nontechnical way. Nevertheless, it tries to present complicated and subtle theoretical and conceptual issues from a range of fields. Occasionally matters become a bit hairy in a way that requires slow and reflective reading, which may be anathema to some. The somewhat more difficult parts are presented in boxes. Everything in the introductory part is used later. My advice: read every part as far as you can, but do keep in mind that skipping to the next chapter (and hopefully coming back at some later stage) is an option. It is not a detective book. You can go back and forth and, sometimes, just move on and never come back. As long as you get the gist of things, I'll be pleased.

Happy reading!

This is a section of [doi:10.7551/mitpress/15932.001.0001](https://doi.org/10.7551/mitpress/15932.001.0001)

How Deeply Human Is Language?

Chomsky, the Brain, and the AI Fantasy

By: Yosef Grodzinsky

Citation:

How Deeply Human Is Language?: Chomsky, the Brain, and the AI Fantasy

By: Yosef Grodzinsky

DOI: 10.7551/mitpress/15932.001.0001

ISBN (electronic): 9780262052023

Publisher: The MIT Press

Published: 2026



The MIT Press

The MIT Press
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139
mitpress.mit.edu

© 2026 Massachusetts Institute of Technology

All rights reserved. No part of this book may be used to train artificial intelligence systems or reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data is available.

ISBN: 978-0-262-05200-9

EU Authorised Representative: Easy Access System Europe, Mustamäe tee 50, 10621 Tallinn, Estonia | Email: gpsr.requests@easproject.com