

Epilogue: Shall We Work Together?

*“Other maps are such shapes, with their islands and capes!
But we’ve our brave Captain to thank:”
(So the crew would protest) “that he’s bought us the best—
A perfect and absolute blank!”*

*This was charming, no doubt; but they shortly found out
That the Captain they trusted so well
Had only one notion for crossing the ocean
And that was to tingle his bell.*

—Lewis Carroll, *The Hunting of the Snark*, 1876

I may not have gone where I intended to go, but I think I have ended up where I intended to be.

—Douglas Adams

An AI Fantasy?

This book’s subtitle is *Chomsky, Our Brain, and the AI Fantasy*. We talked about linguistic theory, inspired by Chomsky; about our linguistic brain; and about AI in its current incarnation—but no fantasy has been mentioned, at least not explicitly.

What’s a fantasy? ChatGPT’s answer (apparently based on *Merriam-Webster*): “Something imagined or created in the mind, typically involving elements that are impossible or unlikely to occur in reality.” Do the arguments in this book support the last, somewhat provocative, part of its subtitle? I believe they do. The fantasy I described is the thought that LLMs alone can provide a true model of human speakers’ linguistic abilities and

brain activity, which I believe won't happen. For this ambition to come true, LLMs would need to pass a much higher bar than the one set by Alan Turing through his test: not only would they need to mimic human language behavior successfully, but also, they would need to constitute *the* true model of the neural networks that underlies linguistic capacity. I have given you good reasons to believe that without incorporating linguistic theory, this is pure fantasy. Let me quickly summarize the arguments, in text and in a table.

The arguments were of three types. *Empirically*, I provided multiple facts—linguistic, anatomical, and neurolinguistic—that suggest that the proposed models, while creative and interesting, must incorporate linguistic theory, or else they remain off the mark as they do not fit the rich available experimental record (which they simply ignore, for the most part). *Conceptually*, I have argued that these models are unconstrained and are moreover opaque; that is, on their own, they are too large to be interpretable. A modular design that responds to grammatical and neurological needs will likely enhance our understanding, without which, we have no good story to tell. Finally, from an *engineering* point of view, I hinted that these models (whose authors keep citing Moore's Law concerning the growth rate of computational power) seem overly muscular. The rarity of engineering efforts to constrain them threatens to lead to energy-hungry machines that kill flies with cannonballs. As theories, none of the current ones is parsimonious. Luckily, neurolinguistics has an alternative scientific agenda, which I presented briefly in chapter 7.

Here are some brief bullet points of the story as it unfolded:

- TGG delivers a scientific theory that describes the knowledge humans possess about their language and that is consistent with neurolinguistic results. The implementability (and runnability) of this class of theories is far from impressive, and its proponents may have been negligent in seeking dialogue with the neural network community. However, we currently have an impressive class of linguistic theories that are transparent and truly explanatory.
- LLMs are task-oriented language technologies, having produced impressive simulations of a speaker/hearer that are runnable on a computer. However, they fail on a surprisingly broad range of psycho- and neurolinguistic tests, they are nontransparent, and they do not deliver a biologically or conceptually feasible (let alone constrained) scientific explanation of the human linguistic ability and its neural underpinnings.

- Neuroanatomy and physiology have provided detailed perspectives about the underlying biology of language and speech. A rich record indicates that the linguistic brain, while massively interconnected, is also highly modular, in a manner that is sensitive to linguistic generalizations. Anatomical linguistic modules, or “brain areas that compute linguistic functions,” seem to be structured in terms of the microscopic properties of the neural tissue that supports them.
- Bridging the gap between anatomical structure and cognitive function is now almost routinely done in vision neuroscience but is still rare in neurolinguistics. This can and should change.

I have addressed a variety of issues, on which the two approaches agree and disagree. The main disagreements, advantages, and deficiencies are summarized in table E.1.

The Necessity of Working Together

Language, it appears, is very deeply human, presenting humanity with a vast array of problems and puzzles. This leads to a clear conclusion: *we must work together!* Linguists should indicate where to look, how to construct benchmarks, and how to structure the functions of our model; neurolinguists should tell us how to modularize our network and how to probe humans and brains; and network experts should build the networks and provide mechanistic explanations for the linguistic generalizations. What could be more natural than this collaborative effort?

Without linguists, claims about the machines’ performance often end up being embarrassingly unfounded; without network specialists, they are not even built; and without neurolinguists, no coherent perspective about brain function can be developed.

Before We Say Goodbye: A Learning Success and a Learning Failure

If your brain analyzed a sentence successfully, then your next encounter with the same type of sentence (e.g., a question, an active declarative sentence, a conditional sentence) would be faster and less taxing, even if the words were different. This has been demonstrated time and again, and it has been dubbed *syntactic priming*.¹ It amounts to a demonstration that

Table E.1

Issue	LLMs	TGG
1. Innateness. Does language have an innate knowledge component?	NO	YES
2. Child development. Does the system solely depend on input to get to its adult steady state? What are the learning methods?	YES Input-dependent, sensitive to quantitative aspects of the input	NO Little sensitivity to quantitative aspects of the input
3. Competence versus performance. Should the theory make this distinction?	NO	YES
4. Constraints. Is the model or theory constrained, or can it grow indefinitely?	NOT really. The few constraints one finds are imposed by hardware limitations	YES, there are many constraints, built into the theory, enhancing explanatory adequacy
5. Neural description. What is the right level of description? Does individual variation matter?	Neural networks/connectivity/layers/weights, simulations	Functional anatomy at the level of microscopic anatomical parcellation; causally modeled connectivity
6. Brain localizers. What method is used to parcellate the brain into its functional parts?	Electrodes; functional ROIs (regions of interest) based on gross distinctions	Anatomic ROIs, based on fine anatomy, mapping with cytoarchitecture
7. Performance tests. What is the right benchmark? How to build a general evaluation method for the system?	Turing test: “imitation game,” “creativity”; gross localizers; behavior/response correlations in next word prediction tasks	Comprehension, judgment, and verification tasks on fine theoretically motivated linguistic contrasts
8. Language evaluation. What are the inventories/test materials?	Language “in the wild,” structured tests (BLiMP, CoLA); lab-specific localizers	Multiple theory-driven choices
9. Runnability. Can the theory be run on a machine?	YES	NO
10. Transparency. Can a failure be attributed to a specific theoretical construct or component?	NO	YES
11. Interpretability. Do we understand why the system behaves as it does?	NO	YES

your brain learned a generalization—that two sentences that are built out of different words but that share abstract structure are similar.²

LLMs do not analyze syntactic structure, but they are said to be fast learners. To their proponents, their ability to generalize from examples indicates amazing learning abilities.³ So here is a new linguistic way to assess this ability, whose existence underscores the need for collaboration between our communities: if the LLM is a fast learner, then once it is taught the properties of a particular structural relation, it should easily generalize to new examples. Pursuing this logic, I set up a two step test of ambiguity detection learning—a task that humans can perform easily—and conducted it with ChatGPT-4o.

Step one tested the machine's ability to detect a basic structural ambiguity discussed in chapter 1. I presented it with an ambiguous sentence: *The first police commissioner was a woman.* (Structure₁: *the [first [police commissioner]]*. Meaning₁: the first commissioner of some police force in the common ground was a woman. Structure₂: *the [[first police] commissioner]*. Meaning₂: the commissioner of the first police force in the common ground was a woman.) Instead of asking whether the sentence was ambiguous, I requested that the machine paraphrase it; the machine returned a single unambiguous sentence, the same as the original sentence but with two words substituted: *The initial police commissioner was female.* (Later, I tried other example sentences, and at times, it was able to identify the ambiguity, although paraphrasing was more difficult.) I replied that my first sentence was ambiguous; ChatGPT agreed to be corrected (as it almost always does, apparently hand-programmed to refrain from challenging its interlocutor), “admitting” that it had previously failed to detect the ambiguity; after minimal prompting, it successfully expressed both meanings. Learning seemed to have occurred.

The critical part was step two, in which I tested the machine's ability to generalize. If the lesson had been learned, I reasoned, and if the machine can generalize, then it would quickly identify the ambiguity in a new sentence that has the same structural properties. This is a true test of generalization ability and a task easy enough for any reasonable human. I therefore asked the machine to paraphrase a close relative of the previous sentence: *The first army commander was six foot tall.* Most humans would succeed in detecting the ambiguity here, especially if asked in the immediate context of the previous instance of structural ambiguity. This is generalization. To

my surprise, I must admit, ChatGPT-4o failed to generalize. Learning had not taken place.

I encourage you to play and invent such games, and I hope you will be amused. More importantly, I sincerely hope that a more serious lesson is learned here: such failures can only be detected once linguistic considerations are invoked (and these are rather minimal). And yet, these considerations are never invoked in the discussion of LLMs' language learning. This type of two step test can be cooked up by anyone with a fairly basic background in linguistics. As LLMs are not equipped with a method to generate predictions and tests of the type described, there is no way to discover these machines' limitations (or virtues, for that matter). It is rather shocking that after a decade of research on learning in LLMs and after much talk about "scientific revolutions" and "paradigm shifts," such simple tests have not been conducted. May it serve as yet another argument for the acute necessity of a collaborative effort between the disciplines.

What to Do: A Sketch of a Recipe

In the preceding chapters, I sketched an ongoing research program that aims at discovering a mosaic of microfunctions and their microanatomical correlates. I gave some hints that suggest the validity of this approach, and I pointed to similarities between it and leading approaches to the study of brain mechanisms for vision. Pursuing this line of reasoning would not only be collaborative in a manner that would cross disciplines and let groups with varied interests teach each other what they know about the field, it would also shift the direction of research from a study of generic networks to investigations of neurologically supported micromodular local networks. Thus, starting off with modules and constraints on their operation, studying the cytoarchitecture of the supporting tissue, and using all this for the construction of small networks is to my mind the right way to achieve progress working together. Even engineers would gain from it, as linguistic knowledge would act as a power saver, helping them trim redundant processes. This is not a fantasy: it's a concrete research program, which would lead to major progress in the study of neurolinguistics and neurocognition.

Noam Chomsky repeatedly said that "linguistic theory is part of psychology, ultimately biology." Adopting the strategy I have sketched would hopefully bring this research program closer to reality.

An Afterthought: AI and MI, October 8, 2023

The Human-Machine team enables a new potential to use AI to build a “smart area” for protecting borders. The ability to connect various bits around the border through machine-learning can help control these areas and protect the borders.

—Brigadier General Y. S. (Yossi Sariel), Commander, Unit 8200, Israel Defense Forces, *The Human-Machine Team*, 2021

On October 8, 2023, when it became clear that the line with Gaza had collapsed, I knew how this book would end. The Gaza Strip, spanning altogether 141 square miles (population: 2.1 million), has for years been thoroughly monitored by eavesdropping devices, camera-equipped drones and balloons, and much else. The Israeli intelligence community, especially the Military Intelligence (MI) section of the Israel Defense Forces, was said to have divided the Strip into data-generating tiny squares, fed into artificial intelligence (AI)-guided devices. Signals intelligence technologies were developed by the famed Unit 8200, which has long attracted the best and the brightest young Israeli minds, becoming the pride of the tech nation and the largest Israel Defense Forces unit. Scores of this unit’s graduates (some of whom would become my students) were hired at top positions in the AI industry worldwide; its commander preached a “Border with ‘AI Walls’”;⁴ and its top brass bragged that “our computers can indicate whether the information is valuable, as well as its value level.”⁵

Every Israeli knew that the borders were watched and secured; that Unit 8200 was there, tirelessly analyzing instreaming big data. The press assured us that everything was fed into superfast computers and chewed up by LLMs. We were also told that, aware of past failures, the MI continually consulted with smart devil’s advocates, who kept questioning basic assumptions and left no stone unturned.

Everyone adored them and slept well, and then, boom! On Saturday, October 7, 2023, this whole creation goes down the drain in less than a day: a murderous attack takes the start-up nation by complete surprise. Hamas staged a well-planned attack, felled the border fences in minutes, and its men murdered, mutilated, and violated some 1,200 innocent civilians, took hostages, and burned their homes. In retaliation, Israel declared a war on Hamas that has since killed dozens of thousands of Gazans, among them almost 20,000 children, wounded even more and has left huge numbers hungry,

wasted, destitute, homeless, and displaced, with no end, solution, or hope in sight.

How was Hamas able to cook up such a surprise? Where were the MI and its Unit 8200? How could the MI's allegedly supreme capacity to eavesdrop and its unmatched AI abilities to predict the enemy's next step miss a huge plan in which thousands were involved?

Uri Bar-Joseph, the renowned military historian of the surprise Arab attack on Israel in 1973,⁶ gloomily described in November 2023:

the tendency in the Military Intelligence (MI) to take lessons of the past lightly, and rely almost exclusively on technological alerts. . . . I saw this tendency myself less than two months ago. In an event marking fifty years for the intelligence failure in . . . 1973, I presented to MI's senior officers my explanation for the roots of this failure. I said that first and foremost, it was the personalities of some past senior officers who stuck to the prevailing conception to the very last minute, despite the fact that all the information in front of them was screaming "war!" A subsequent lecture at that event was about an experiment, in which available pre-war data were fed to an AI machine, in order to examine whether AI can substitute human thinking. The discussion period was mostly focused on the machine's abilities to identify threats. The psychology of evaluation failures was not of interest to the MI evaluators. As I left the discussion, it was clear to me that the lessons from 1973 were not learned. Not for a moment did I imagine that it would become so painfully and shamefully clear at this speed.⁷

By early December, it was becoming clear that heaps of evidence had been sitting right at the MI's door. Hard questions about basic assumptions were not asked, as the burden was shifted to LLMs, which did not converge on the right evaluations. On-site human observers continuously sent clear warnings, but neither machine nor high command listened.⁸ What had been touted as the greatest military intelligence service ever fell victim to a gigantic deceit by Hamas's guerrilla army, in the face of heaps of evidence for an imminent attack. All warnings went unnoticed.⁹

Strategists and experts will spend years discussing this tragic surprise, but one thing already seems clear (in fact, almost a consensus among experts): the MI and its satellite organizations became addicted to high tech, relied on developers, and failed to think hard about the right characterization of the problems that these technologies were expected to solve. They believed that their devices were so powerful that they could offer a solution to anything and spent insufficient amounts of thought on whether they were choosing the right problems to solve. Thus, seeking prediction without clear definitions of the problems that lie ahead, ignoring large datasets and relevant

biases in the input data, and failing to evaluate correctly or at least consider alternatives seem to have been central factors in Israel's worst intelligence failure ever.

Why am I dwelling on intelligence at war in a book on language? Because this (still ongoing) Middle Eastern tragedy is highly relevant to the debate on the human linguistic ability, to which this book is dedicated. A neurolinguist living in Tel Aviv, I find it difficult to miss the similarities between the atmosphere within the Israeli MI prior to October 7 and certain prevailing sentiments in the AI community that models natural language and its neural bases.

I love technology, and I love science—incontrovertibly among the best products of human culture. But overreliance on technology (and to some extent on science) may lead to complacency. Hence the fate of the AI wall story, a fantasy that everyone loved to believe, until it collapsed, as border fences fell in a matter of minutes. Tragically, the largest and smartest unit of one of the strongest military organizations on the globe failed to predict a surprise ground attack by guerrillas riding in low tech white Toyota Hilux pickups.

In addition to complacency, overreliance on technology threatens to lead science astray. On the face of it, the practices of LLMs' proponents who build language technologies are strikingly similar to those of the Israeli MI. As Rich Sutton puts it, more technology, more big data, more memory, and more computation would make unsolved problems go away fast. This belief has become widespread. Indeed, when my AI-heavy neuroscientist colleagues are presented with problems that ChatGPT has had (something I have been showing them for a while), they respond, unwaveringly, "if ChatGPT would only train properly, he/she/it will eventually get it." Their belief in the power of AI appears to run deep.

And yet, a mere year and nine months ago (we are now in July 2025), a critique of the Israeli MI and its AI capacities would have been dismissed as heresy. The Israeli technological disaster, I believe, should echo far and wide. This book, aimed at both lay readers and experts, attempts to spell out problems with the use of language technologies in science and elsewhere. In it, I have tried to consider alternatives, in the hope of learning an important lesson for a better future. Presently, I approach stories about great technologies and, to some extent, about science with awe but also with suspicion and fear. And in the case dearest to my heart—human language—I see great successes but also serious problems and perils, and I wish that readers would keep this in mind.

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