

Claremont Colleges

Scholarship @ Claremont

CMC Senior Theses

CMC Student Scholarship

2022

Language in the Age of Artificial Intelligence

Axel Ahdritz

Follow this and additional works at: https://scholarship.claremont.edu/cmc_theses



Part of the [Digital Humanities Commons](#), and the [Other English Language and Literature Commons](#)

Recommended Citation

Ahdritz, Axel, "Language in the Age of Artificial Intelligence" (2022). *CMC Senior Theses*. 3175.
https://scholarship.claremont.edu/cmc_theses/3175

This Open Access Senior Thesis is brought to you by Scholarship@Claremont. It has been accepted for inclusion in this collection by an authorized administrator. For more information, please contact scholarship@cuc.claremont.edu.

Claremont McKenna College

Language in the Age of Artificial Intelligence

submitted to
Professor Leland de la Durantaye
and
Professor Mike Izbicki

by
Axel Sverker Ahdritz

for
Senior Thesis
Fall 2022
December 5, 2022

Abstract

AI language models can now produce text that is indistinguishable from our own, forcing us into a confrontation with the romantic assumptions underlying ‘natural language’ in the West. In this thesis, I will conduct a genealogy of the ‘natural’ and ‘artificial’ qualities of language through the literary, philosophical, and mathematical texts in which our ideas of authorship are premised. My hope is that this discussion will deepen our understanding of the language produced by AI models, answer why we feel compelled to anthropomorphize these machines, and situate readers in the reality of our present linguistic moment.

Acknowledgements

First and foremost, I would like to thank my teachers. Wesley Phillipson, I am forever indebted to you for teaching me how to write. You taught me to pay attention to language, and in so doing, you are truly the generative force of this thesis. Leland de la Durantaye, you gave me exactly the right books, exactly when I needed them, whether it be Borges or Janes, Auerbach or Joyce. I cannot tell you the extent to which you have fostered the growth of my soul. Mike Izbicki, I will be forever grateful to you for teaching me practically everything that I know about data science, and for your continual support and advice over the past many years. And Joanne Ahola, you walked me through the indefinite and impenetrable parts of my mind and taught me how to see.

Sanjana Gauri Bhatnagar, you have always been the moon that guides me, but I literally couldn’t have written this thesis without you. You REPEATEDLY saved me from my pits of despair and REPEATEDLY dropped your entire life to aid me. Your editing skills are sublime, and the power of your mind is immanent in every word you speak. To my roommates, Emily and Shraya, I thank you for welcoming me into your hearts and for building the community that the kid inside me has always sought. And to my old roommates, Calder and Simon, I owe you my deepest gratitude for calling my bullshit throughout the years and for loving me despite it all.

Lastly, I would like to thank my family for supporting my erratic decision to venture out to California. Mamma och Pappa, det finns ingen som förstår mig som ni gör. Jag älskar er mera än mina ord kan uttrycka. Erik, I regret not being present for more of your life while growing up; you are both the brightest and the most determined kid I know. And Gustaf, I always have you in my mind, checking my statements, and taming my words. You are the best editor I have known, even though you have never explicitly read anything I have written. All I need to do is imagine what you would think, and the answer appears before me. To you all, thank you, thank you, thank you.

Table of Contents

<i>Chapter 1: Background</i>	1
The Natural and Artificial Dimensions of Language	1
The Exponential Archive.....	10
The Language Model.....	17
<i>Chapter 2: The Book of Life</i>	32
The Anthropomorphized Machine	32
The God-Author.....	38
The Romantic Author	47
The Language Model under the Natural Paradigm	58
The Divinatory Origins of Artificial Systems of Interpretation	63
<i>Chapter 3: The Book of Death</i>	68
The Tomb of Language	68
The Formal Text	80
The Death of the Author.....	92
<i>Conclusion</i>	104
<i>Works Cited</i>	106

Chapter 1: Background

The Natural and Artificial Dimensions of Language

Several months ago, while parsing through the diaries of August Strindberg¹ in the dim light of the Library, I found myself immobilized in the space between what my pen had just scratched onto my notebook—*jag är*—and the word which was to follow—*osäker*. Something about the physical shape of *jag är* ('I am') seemed a little off. Holding my notebook up to the light, I looked over each dotted vowel and verified that the consonant loops jugged out in the right direction. Unsatisfied yet convinced of my own lucidity, I allowed my pen to trace out the next word in the sequence—*jag är osäker* ('I am unsure'). I paused once again. What was I unsure about? Grumbling, I turned back to the page in Strindberg's volume and diligently performed the act of 'reading'. With empty eyes, I scanned its lines from left to right. I stopped at periods. I felt the shape of each word in my mouth. I listened to the vowels and the consonants. Upon reaching the end of the page, I began to think. But my mind grasped at vacant space: I hadn't comprehended a thing. I tried again. Nothing. The black dotted marks on the page were no more intelligible than cracks on the surface of a rock. I laughed uneasily at the senseless scrawls before me, for I had forgotten how to read.

However brief my lapse of understanding was, these moments of defamiliarization allow us to take conscious stock of the physicality on which our daily experiences rest, to regress back to the early stages of childhood in which things were exactly what they were, nothing more and nothing less. Illuminated by our conscious apprehension, the most obvious things—forgotten for the sake of fluidity or function—can suddenly appear before us, unfamiliar and terrifying. If we

¹ Strindberg 8, for those that are curious.

think too hard about *what we are actually doing* while driving a car down the highway, recognizing the explosive metal box surrounding us and the speed at which we shoot down the concrete roads, I think the terror would immobilize us. The same suspension of physical reality must be undertaken by the concert pianist, who, in the ecstatic heights of a sonata, must resist focusing on the autonomy of her individual fingers. The steady click of a clock will grind to a halt under the steady gaze of those who wait, and a specific word, repeated over and over, will quickly deteriorate into meaningless chatter.

Sometime after the early years of childhood, we learn to forget the physical form of language. We learn that the words themselves do not matter, only the concepts, referents, and spirits dancing behind them. Like skilled aerialists leaping from concept to concept, our minds willfully forget the arbitrary signs on which said concepts rest (a state often called ‘fluency’). Ideally conceived as a formless mirror—a mimetic substance that reproduces nature and divinity—the physical form of language becomes an invisible intermediary that links us to the spirit of the author² and to the world they exist within. Our ideas of language are structured by our first relationships with our beloved caregivers, connecting with our mothers and the culture we will spend the rest of our lives within. In hearing the voice of a loved one, do we not feel their presence, the embodiment of everything that is them? So great is the mimetic illusion that from the moment we are born, we are embedded within the linguistic structure of a *name*, onto which all our lives are reflected. It suffices to use a name to bring a person into another’s mind. Coevolving alongside humankind for the past two hundred thousand years, language does not merely represent life—it is often afforded a life of its own. We speak of *living* languages as compared to those that are *dead* (e.g., Italian vs. Latin). But to participate in language, to partake of the collective communion, we

² I will be using the term ‘author’ to designate any subject that produces language (be it in speech, or in writing).

must transubstantiate the form into the spirit. We must take the text only for what it *means* and discard the rest.

This romanticized conception of language as a natural, ‘living spirit’ often suppresses the artificial and conventional forms on which it is premised. Constructed from mutable objects—from the structural architecture of a word, down to the microscopic layer of ink or pixels in which it is rendered—written language is first and foremost a physical artifice. Sometimes it takes another language to see it:

السلام عليكم³

For those that do not speak Arabic, the physical form before you will refuse your attempts to enter. Moving your eyes across the screen, your eyes might locate the break between its words and letters, but already, you have failed. Arabic is read right to left. We are hopeless without the normative rules and practices on which our understanding is premised, institutionalized through spelling, vocabulary, syntax, grammar, styles, registers and more. Our sentences are constructed like bridges, their conjunctions and particles are hammered in like bolts, fixing together a set of parts predefined by the system.

Even *verbalized* utterances have a physical shape in the throat. Phonemes constrict our tongues and our vocal cords just in time for our breath to vibrate through. When I say, “*sbbh... listen*” I am commanding you to enter a state of silent awareness. But to do so, I must pull my tongue to the back of my throat, partially close my teeth together, move countless muscles in my face to fashion my lips into an ‘o’, hiss like a snake, and then pause for a predetermined length of

³ Arabic, “Peace be upon you”

time... and that is before I reach the second word.⁴ The signs are predefined, and so is the imperative formulation of ‘listen’—the verb without a subject—in which the brunt of my statement lies. All of this is picked up on your end through an equally complex decoding network. Only then, after an elaborate series of actions drilled into both of us since childhood, is ‘silence’ achieved. It is undeniable that the artificial system in which my utterance is encoded is performing a lion share of the work that we colloquially call ‘meaning’.

Equal and opposite to the natural, intentional, and spiritual connotations of language is its brute, material existence. Few recognize this fact more than the writer or poet. Engaging in forgery, insincerity, and falseness, they reproduce copies of the natural world. They assume divine authority—bestowed by a god or the Muse—and have the audacity to call their fiction truer than life. They (usually) cover up their tracks. Nonetheless, the esteemed poet must still adjust to the physical limitations of the system. They are sly servants of grammar; their stanzas are steered by rhyme and rhythm as much as they are by ‘reason’; they are tyrannized by semantics, computed by meter, and encompassed by myth. For their work, they have been called saints and devils. Plato cast them as demiurges of a lesser reality; Heidegger heralded their language as creating the very world in which we live. They spend their lives erecting empty houses and haunt them for eternity.

There is an air of the grotesque, the sacrilegious to this mad artificing. Creating idolatrous monuments to their own self-perpetuation, the writer’s consideration of the physical object of language defamiliarizes our ideas of where and how meaning is located. Their characters are not the authors of their own words, and to some extent, relying on the system of language to construct

⁴ Or, in the case of the words before you: order a series of mental representations of words and sounds into their component letters and then translate them into rapid movements of my fingers, which are then encoded as electronic pulses in the CPU, and reconstituted through predetermined routes through the operating system, et cetera

their meanings, neither are they. But this must be kept a secret, lest the reader's illusion is broken, and their work be denounced as contrived and artificial.

Though the system of language has a large hand in controlling the meanings we can create—from low-level grammatical structures and vocabularies, to higher-level tropes, themes and genres—there is a pervasive attempt to obscure these structures and to resist the artificialization and objectification of language. We believe in the intention of a spirit behind language, for how else could we understand it? Words, sentences, books all possess no meaning by themselves. They are ontologically dead—mere squiggles on a page, governed by a set of artificial rules, laws, and patterns. In order to preserve meaning, preserve the spirit of language, we must rely on the idea of the author to fasten our text to a particular time, place, and mode of discourse. The author—be it through speech, writing, or any other signifying capacity—is the guiding principle by which we determine (or imagine) life within language. We orient language in discourse by considering the identity of the individual that created it, the context in which it was said, to whom they are speaking, alongside a host of other inferences.

It is often of existential importance to preserve the spirit of the text. In the Abrahamic context, God's divine utterance on the first day of creation, "Let there be light", rendered all life possible (*King James Bible*, Genesis 1:3). Here, language does not merely mirror Nature—it created it. The language of God and of his prophets, embodied in Judaic, Christian, and Islamic law, demands its own recitation every day of the year, its body kept warm in the throats of billions. Indeed, "In the beginning was the Word, and the Word was with God, and the Word was God" (John 1:1). The spirit of God is literally manifest *as* the language attributed to him, not within it, nor outside of it. To consider these texts as man-made artifices is a sin punishable by death and an eternity in hell.

We also see the authorial principle at work while browsing text online. Any word we read online is explicitly attached to a name and a context. The tweet is not simply composed of its text. Its paratextual components (borrowing from Gérard Genette) are *literally* embedded within the text in the form of metadata, specifying the relevant information of the tweeter (their photo, their bio, their verified status), what they were responding to or retweeting from, who they referenced or hashtagged, their location on the earth and more. Wherever this tweet travels, this information follows. We refuse to allow text to become unidentifiable—a mere object—for without its context, the entire textual landscape would devolve into meaningless chaos (and the economics of targeted advertisement underlying the entire internet industry would collapse). This resistance towards the objectification of text will be discussed at length in the first chapter of this thesis, but for the time being, let me mention that this resistance is not limited to the West or even strictly to religious contexts. Whether in the Vedic texts or in the words of Confucius, in the divinatory power of the Yijing or that of the Odu Ifá, in every literary text, article, and constitution, every speech, obituary, or proclamation of love, our connection to the natural world, our ancestors, and each other are premised on the assumption of life humming within the intentional sign.

Language is forever at war with its physical reality, a conflict between the natural and the artificial, the spiritual and the material, the intentional and the unintentional, *the living and the dead*. The outcome is certain and grim—for the dead will win. Time will awaken from its slow slumber and extinguish us. With the loss of authors and interpreters, the body of language will have no spirit to animate it. Our symbols will be subsumed by their brute physical existence; a lonely monument, a metal plaque, buried deep in the fossil record, dreaming of nothing. For some, this is the purest of freedoms. Roland Barthes in declaring “The Death of the Author” in the mid-20th century makes this belief clear: “the book itself is only a tissue of signs, a lost, infinitely

remote imitation....Refusing to assign to the text... an ultimate meaning, liberates an activity which we might call counter-theological, properly revolutionary, for to refuse to arrest meaning is finally to refuse God and his hypostases, reason, science, the law” (Barthes, “The Death of the Author” 1271). However, I do not know if Barthes himself has fully considered the sheer proliferation of meaning caused by the removal of the authorial spirit.

Once flattened into its physical form, once purged of its spiritual dimensions, the text is left as a mutable, meaningless material, at the whim of our combinatorial fancy. Nothing captures the sheer mathematical terror of this state better than “The Library of Babel” by Jorge Luis Borges (a man I both curse and revere). In this short story, Borges conceives of the universe as a library “composed of an indefinite, perhaps infinite number of hexagonal galleries” whose books contain every possible combination of the letters of the alphabet, spaces, and punctuation marks (Borges 113). Humans (now librarians) are condemned to wander its labyrinthian halls for eternity, crowded by endless tomes of senseless cacophony—pages and pages of the word *abdcdfafngot* over and over, countless books consisting of different combinations of commas and periods, single sentences with every word in the Spanish dictionary in no particular order. Meaning must be found in this senseless life, so the librarians attribute its text to the gods, spending their lives searching the books for any modicum of coherency. For somewhere in its numberless shelves there will exist “the detailed history of the future, the autobiographies of the archangels... the true story of your death” and critically, “the faithful catalog of the Library”: an elucidation of the linguistic system that would allow them to decode their existence (Borges 115). Those that repudiate the quest to find meaning in such a Library are deemed “infidels”:

One blasphemous sect proposed that the searches be discontinued and that all men shuffle letters and symbols until those canonical books, through some improbable stroke of chance, had been constructed. The authorities were forced to issue strict orders. The sect disappeared, but in my childhood I have seen old men who for

long periods would hide in the latrines with metal disks and a forbidden dice cup, feebly mimicking the divine disorder (Borges 116).

Despite their recognition of the true operational structure of the Library, it is of existential importance to the librarians—those religious fanatics of the sign—that the texts are not the product of random rearrangements of orthographic signs. The structure of their civilization depends on an idea of an intentional God-Author behind the text.

But, for every ‘true catalog’ there are “thousands and thousands of false catalogs” and even “a proof of the falsity of the true catalog” (Borges 115). Without a method of attributing meaning to the texts, the Library will ultimately “affirm all things, deny all things, and confound and confuse all things, like some mad and hallucinating deity” (Borges 117). For Borges, this does not render the texts *meaningless*, however, pointing out that the phrases of letters

at first apparently incoherent, are undoubtedly susceptible to cryptographic or allegorical ‘reading’; that reading, that justification of the words’ order and existence, is itself verbal and, *ex hypothesi*, already contained somewhere in the Library. There is no combination of characters one can make—dhcmrlchtdj, for example—that the divine Library has not foreseen and that in one or more of its secret tongues does not hide a terrible significance. There is no syllable one can speak that is not filled with tenderness and terror, that is not, in one of those languages, the mighty name of a god. To speak is to commit tautologies (Borges 117).

If every utterance were deprived of authorial intention, if language were to be flattened to mere combinations of orthographic symbols, language would become impossible and blasphemous. Even to speak would be to “commit tautologies”, for language would signify all things (by signifying nothing). Like attempting to read “the chaotic lines of the palm of one’s hand” (Borges 114) or the “awful Chaldee” on the sperm whale’s brow (Melville, *Moby-Dick* 263), any attempt to find meaning would be commensurate with understanding God’s creation.

For Borges, *language is the Library*. For humankind, *the Library is the universe*. Solely humans are trapped in this physical artifice. They are the “only species” in Borges’ world (Borges 118).

This conception of language is consistent outside of Borges' text. We embed all we know and perceive into our linguistic system. We know a 'bear' because we have named it (regardless of whether we have seen one). We take notice of the 'proletariat' only after this word is situated in the system of communist thought. Our world is constructed through the artificial system we have created, and it enables and limits everything we are able to understand. Subsumed within this monstrous artifice, our civilization depends on resisting the artificialization of the linguistic sign. Borges foresees the final victory of the dead, physicalized existence of our language at the end of his story. While the human species "teeters at the verge of extinction... the Library—enlightened, solitary, infinite, perfectly unmoving, armed with precious volumes, pointless, incorruptible, and secret—will endure" (Borges 118).

This thesis is principally concerned with how the polarity between the *natural and artificial*, duplicated in the binaries of the *spiritual and the material*, the *intentional and the unintentional*, the *living and the dead*, are in constant tension within language, and how this tension can help us understand our relationship to the literary past, the burgeoning textual archives of the present, and the cybernetic future of artificial intelligence systems. In this section, I discussed how the artificial components of language are hidden from our view—as much by familiarity as by our spiritual conceptions—and the inordinate role these artifices hold in delivering meaning in discourse. Then, I briefly introduced the author—the spirit that hums within the linguistic sign—and the meaning-fixing role they play in our language. Finally, I expounded upon humankind's resistance towards the artificialization of language. Before concluding this first chapter, we must examine the proliferation of our textual archives in the technological age, and subsequently, the ungodly machines that this proliferation has produced.

The Exponential Archive

Before the advent of writing, before movable type revolutionized the distance our words could travel, before the hard drive and the database, much of our spoken language—thrown out into the air—simply disappeared after its use. Only our memory could safeguard words from the passage of time. History was kept in beneath the greying hairs of elders, legends in the throats of bards and griots, and proverbs in the knowing glances between parents and children. Memory is fluid, like us, which means that language could not be fixed in time or place. Our words diversified as we dispersed across the globe, encountering new animals, materials, and spirits in our environment. Stories were liable to the creative freedom of the teller and were premised on the traumas and victories of a community. All that was not passed down—in one way or another—would die alongside the human mind that carried it. Apart from the symbolic functions of early tools and pictorial representations, language was entirely encompassed by our living biological network. It is in this setting that language lead its most natural life.

Despite the lack of an external repository of our language, the artificial structures of language still held archival power. By speaking, we subtly reinforce patterns or alter existing ones in our language. By producing neologisms, we externalize our language into the larger structure, even if the etymology of the word becomes lost to time. Particularly apt turns of phrase and metaphors are integrated as idioms.⁵ Some infinitesimal portion of every individual's language is archived in the system, an imprint that is picked up unknowingly by future generations of speakers. The system of language itself stores our memory.⁶ This doesn't mean that every individual has equal effect on the system. Which stories live on, and which die out in human culture are often a

⁵ We “grasp” ideas and concepts—a metaphor of the “mind as a hand” that we do not think of in its use.

⁶ Linguistics allows us to trace the movement of vast groups of people throughout history based on this embedded memory.

function of their mnemonic power, making Narrative, Melody, Rhythm, and Rhyme some of the earliest tools for linguistic preservation. Great tales, carrying the language of our ancestors, are unforgettable. Their archetypal characters and thematic elements are broken down to refurnish countless subsequent stories (think the recurrent figures of the ‘trickster’, ‘troll’ or ‘dragon’ or the thematic arcs of the ‘heroes’ journey’ across cultures worldwide). In setting a narrative to song, a poet plays upon the cyclical patterns of rhythm and our need for melodic resolution. This organizes the story in the memory according to language’s external physical properties, like its syllables and sounds. Proverbs are preserved in language due not only to the wisdom they express but also to their ability to express it succinctly (often with rhyme). The artificial structures of language evidently have major archival power. Nonetheless, any system premised on memory and habituated linguistic use cannot hope to fix language in time. It can only reproduce it again and again, a body that constantly mutates in the mouths of others.

The preeminence of the natural archive of memory was upset upon the advent of the written word. By encoding the sounds and representations of language into a symbolic system governed by a set of artificial rules, meaning could be (hypothetically) stabilized and maintained absent the immediate presence of a speaker and their context. Writing is thus a kind of compression algorithm that reduces the natural dimensions of language such that it can be deterministically decompressed by another schooled in the algorithm. The storage capacity of the memory, subject to natural dissolution, no longer limited what was remembered or forgotten. The mnemonic devices of narrative, metaphor, melody, rhyme, and rhythm were no longer the sole judge of what remained throughout the centuries and the generations. Anything could be encoded into writing, from the most banal food-rationing system of the Mesopotamian kingdoms to the laws of Hammurabi. Freed from the throat, the memory, and the bounds of the local culture in

which it was previously situated, language now acquired a fixed body of its own that was neither dead nor alive, one that was agnostic to the passage of time.

I do not have the knowledge nor the space to expound fully upon the history and impact of written language here. Such a task would take a lifetime to complete. So, let me be extraordinarily brief. Physical records and asynchronous communication allowed for the growth and maintenance of larger and larger cultural bodies, which subsequently required new languages and artificial structures to describe and manage, which further expanded our culture's dependence on these structures. The result: the exponential growth of the archive and the artificialization of language. Before discussing this topic in greater depth, let us take a second to meditate on the vast differences between the written archive and the memory archive.

Though the storage capacity of the memory is nothing compared to the physical limits of space, a material text cannot be stored, maintained, nor accessed in the same manner as our memory. Written language—physicalized in clunky books, clay tablets or what have you—can only be stored in a single place at a time. To speed up information retrieval, text in a library must be organized according to some artificial system of attribution, such that someone can index the collective memory based on the author, publication year, subject-area, et cetera. Without any natural mechanism by which irrelevant texts may be discarded, our textual archives must be organized hierarchically based on their importance. Memory, on the other hand, knows none of these distinctions. As a rhizomatic network of knowledge, our minds can store our memories in a vast and interlinked web without any of the library's codes and systems. From within the same operating system, we can access pertinent information in a matter of milliseconds, formulate them into words, utter them aloud, attach images to them, reexperience them, explore our feelings about them, and discard them during sleep. While the archive of physical text is dead, the archive of our

memory lives. The deficiency of the written archive was captured succinctly by Vannevar Bush in his article for *The Atlantic* at the conclusion of the second world war, “The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships” (Bush). To construct an interactive, indexable, and self-operating archive was humanity’s next task.

The solution was offered by a new encoding system for language: binary. Whereas written words represented the concepts and sounds of spoken language, binary encodings are yet another step removed, representing the physical letters or logograms of written language as a combination of 0s and 1s. Unlike writing, binary is not built to be read by humans. Take these 51 bytes before you:

```
01001001 00100000 01100010 01110101 01110100 00100000 01110000 01110101
01110100 00100000 01110100 01101000 01100001 01110100 00100000 01100010
01110010 01101111 01110111 00100000 01100010 01100101 01100110 01101111
01110010 01100101 00100000 01111001 01101111 01110101 00101110 00100000
01010010 01100101 01100001 01100100 00100000 01101001 01110100 00100000
01101001 01100110 00100000 01111001 01101111 01110101 00100000 01100011
01100001 01101110 001011107
```

Chances are, the words you are reading right now are encoded in a similar fashion. Though they are visually rendered as words before you, this is merely an artifice produced by the word processor or the web-browser you are reading on. The rules governing a character’s encoding and the

⁷ “I but put that brow before you. Read it if you can” (Melville, *Moby-Dick* 263)

processors that are required to display it are far more complex than you would imagine.⁸ The first layer of artificiality is the electricity signals on the RAM or CPU which are either on or off. The second layer is the conceptual leap of associating these states with the numerical representations of 0s and 1s, which itself are defined by the artificial language of mathematics. The third layer is the layer of the unicode decoder—a map that associates strings of these digits with the letters of the alphabet (and the words of every other language). The fourth layer is the word processing program that renders the words on the screen before you in a particular font/format, itself built off mathematical logic (connecting back to the transistors). The last layer of artificiality is the writing system itself, whose physical properties we have already discussed. Only then do we get back to language.

The great majority of our archived language is in this format today, saved as magnetic North-South cues on the hard drive (representing 0s and 1s). The modern SSD (solid-state drive) in your laptop can store anywhere from 0.2 to 4 terabytes of data in a space no larger than the palm of your hand. This means that up 48 copies of the entire library of Alexandria, an estimated half a million scrolls, can now be carried in your back pocket. The mechanization of our linguistic archive into unreadable, microscopic quantities has enabled its exponential growth. Before the introduction of the printing press and moveable type, the number of manuscripts in Medieval Europe numbered in the thousands. By 1500, there were approximately 9 million manuscripts in circulation (“The Medieval Book”). Today, there are 130 million books that have been published since printing was invented 660 years ago according to Google Books. For reference, this means that the total archive of information existing in published books is something like 21 terabytes (if

⁸ The development of the Unicode standard is a feat of astronomical importance for cross-cultural communication online. I encourage you to look it up.

we consider only the information contained in the words and not any metadata, with 300 words per page and an average of 300 pages). It is difficult to pin down the exact amount of text on the internet, but just to give you a sense of it, let's consider Twitter alone (a small fraction of the internet). There are over 500 million tweets produced daily—that is 6 terabytes worth of text every year with an average tweet length of 28 characters (again, if we only consider the text without any metadata or media).⁹ This means that Twitter alone creates more text in 4 years than has existed in books for the past 660 years. As for the text on the internet, one can only imagine. The entirety of the internet archive has doubled every year since 2011. The total size of the internet was around 295 exabytes in 2011 (Stewart). Today, its total size is somewhere around 5 zettabytes, or 5 billion terabytes. If this data were encoded into words, it would require 12 quadrillion books. By next year, this number will double to 24 quadrillion books.

The unprecedented size of our textual archive has transformed the body of our language into something unrecognizable to us. Now a measurable quantity, encoded into mathematical oblivion, our language can be processed, moved, translated, summarized, generated, and transmogrified into pictures without any human involvement. Some might call these processed 'artificial' in and of themselves. But these Natural Language Processing (NLP) processes are premised on every layer of artificialization that has allowed the mutable, constantly shifting qualities of mental language to be transformed into a fixed and heavily controlled system existing external to us, a process of artificialization that began with writing.

The artificialization of language—the frame by which we understand the world around us and communicate it—has heralded the artificialization of the world. Everything is no longer

⁹ https://blog.twitter.com/engineering/en_us/a/2013/new-tweets-per-second-record-and-how

actually what it is anymore, but a symbol in our encoding system. A recent paper on the science of artifacts by Svante Beckman describes this process:

One aspect is that progressing symbolization including both the growth of available information (clearly exponential at least since the 17th century) and the growth of what Popper calls “objective knowledge” which we can approximate as the descriptive content of all sciences. The process of increasing symbolization could also be indirectly measured in terms of the physical storing and transmission capacities of communication systems. Secondly, we find progressive institutionalization—rule-binding—of human behavior individually and collectively. Together these two forms of ideal artificialization mean that the biosphere is covered by the thickening mist of the nousphere—a sphere of meaning (Beckman 63).

Our “nousphere” keeps us swaddled in our own artificial systems, a process that ultimately starts with our symbolization of it. In naming something, one is now establishing a physical quantity to be measured and manipulated. Scientific language in every discipline has allowed the sum of human experience and our natural world to become an interrelated set of symbols and quantities in data sets, one that is no longer premised on what we can see or experience ourselves. We named a “black hole” before we could even observe one. We suffer not from body pains but from “cholesterol”, “melanoma” or “diabetes.” Even the natural world has been reduced to sentimental patches of symbolic ‘nature’. The expansion of our language has purged the gods from the stars, killed the spirits in the waters, and flattened the ineffable and unknowable dimensions of the world. Anything left out of this system of naming is invisible or nonexistent. We must take a picture of a family gathering, or it no longer occurred. We must tweet about that thing we said to our friends, or it will be forgotten. We have rejected the natural state of spoken language—a substance that is ephemeral, mutable, and that disappears upon use—and see it as deficient now. Our lives can no longer live within our mouths and our memories. They must be imprinted into our system of fixed representation, outsourced to the archive.

The Language Model

Some say that ideas can never be destroyed. But words can be misspelled, grammatical rules can be broken, and books can be burned. Once reduced to the status of a physical object in space, indelibly separated from the mind of an author, there is nothing left to preserve the integrity and intention behind language. I can read it or interpret it any way I wish to. I can use your book as a stepping stool or to prop up my shelf. I can take your words as an anagram, reading it only for the first letters of every word, as the Kabbalist did with the Pentateuch. I can splice your words together with the words of a white supremacist and call you racist online. Language, once physical, can be taken for its parts and deliberately misused, an art which some, like the OuLiPo (*Ouvroir de littérature potentielle*), delight in. Many object to these practices, calling them unnatural and artificial. As we discussed in “The Natural and Artificial Dimensions of Language,” it is essential for us to maintain the illusion that language has intrinsic meaning, that the author still exists within the sign. In construing another’s language as a physical object, it seems to us that one is killing the spirit itself.

Even in the age of computers, where language has been transmogrified into numbers and reproduced onto millions of computers around the world, the modern text is still considered to harbor the original spirit of an author, to be uniquely determined by an individual’s intention, to be—for lack of a better word—*alive*. In this blind belief, we fail to understand how the transformation of language into a deterministic encoder-decoder system—first with the written word and then the binary word—has outsourced much of our meanings and our memories into an artificial superstructure that is entirely independent from us, indifferent to our attempts to make something of it. The sheer size of our modern archive, the deterministic and recurrent patterns that dictate our meanings, and the knowledge we have embedded into our linguistic system has

enabled the rise of an author we have never encountered before. An author that challenges our romantic ideals about nature and meaning. An author that dethrones us from the seat of our language.

Enter Artificial Intelligence. Striking fear into the heart of the humanist and eliciting wonder in the mind of the engineer, this name carries more weight today than the ineffable tetragrammaton.¹⁰ But thrown around as it is in sci-fi novels and venture capital pitches, there are few that understand what it is, and how it works.¹¹ In this section, I will explain AI language models in the simplest terms possible, for demystifying their function is key to understanding their role in the developing history of our language. My hope is that anyone, regardless of their background in mathematics or computer science, will be able to roughly understand how AI is able to “read” and “learn” our language by the conclusion of the section.

Let us start by discussing machines in general terms. I will be using ‘machine’ to refer to any apparatus that uses or applies mechanical power through the interrelated function of several component parts to perform a predefined task. I would like to highlight several principles underlying the operation of machines and burn them into your mind. The confluence of advanced humanism, our natural conceptions of language, and increasingly complex machines have produced far too compelling an artifice. We must combat this illusion:

1. *Machines act on physical objects*. No matter how complex a machine may be, they cannot act on ideas or on immeasurable quantities. A sewing machine does not act on the word ‘silk’ or on the idea of ‘silk’—though a skilled human operator might ascribe these higher order functions to the machine. Rather, the machine acts on the physical object of the textile.

¹⁰ The name of God, “YHWH” or “I am that I am”.

¹¹ You might mark me as an overly abstract humanist, thoroughly obsessed with language, but I am also a practitioner of data, spending much of my time marveling at machines.

2. *Machines can only act on specific kinds of physical objects.* The sewing machine raises and lowers a threaded needle. If you input two slabs of wood into the machine and attempt to sew them together, or if you whisper a couple of words into the machine and ask it to weave a sentence, it will undoubtedly fail. The needle was not built to handle the structure of the wood's grain, nor was it endowed with the microphones, the processing power, or the Fourier transformation algorithms (and the like) required to process raw human speech. The physical inputs that a machine can take are often *explicitly* defined at the time of its development.
3. *Machines act on physical objects using a set of explicitly defined rules.* That is, for a machine to process a particular object as its input, they must follow an explicit procedure. In our sewing machine example, these programs are mechanically represented in its circuits. If one pushes the pedal down, one can adjust the speed of the needle because the pedal is explicitly connected to it. Machines always have rules, even if these rules are not hand-made.¹²

Computers have become so powerful and multifaceted today—handling a host of inputs from images to mathematics to language—that we often forget that they too are machines. By the first principle, the transistors in a computer operate only on the physical object of electric binary signals. By the second, they can *only* process this kind of physical input. Any higher order input that is given to a computer must be explicitly converted into this physical form before being operated on. A simple black and white image is not an *image* to the computer, but a grid of pixels in which each square is a binary digit, with 1 representing black and 0 representing white. The

¹² Some might object to this principle, citing the fact that humans can decide to use machines in whatever manner they chose to. For instance, I can hammer in a nail by picking up a sewing machine and dropping it on a nail. In these cases, I object that it is not the machine that is acting, but the human.

computer doesn't look at language as we conceive it, but rather sees strings of binary digits representing letters, punctuation marks, and pictograms—physically represented as electric pulses. By the third principle, computers act on these electric signals using explicitly defined programs. What makes computers seem 'special' is their ability to be programmed by users after its development. Though the hardware itself embodies a certain base program in its circuits, it can run numberless other programs on this basic structure. However, this doesn't mean that the computer is any different from any other machine. The programs running on the hardware are themselves machines, premised on the computations offered by the transistors, taking only explicitly defined inputs, and processing them as per their codified rules. The computer can thus be seen as a kind of telephone switchboard, in which the superficial circuitry of the machine is altered by the direct actions of its users.

What we call 'artificial intelligence' is one such machine running on a computer's hardware. By the first principle, they operate on physical objects, not ideas. By the second, they are built to handle only a specific kind of object as its input. By the third, they process said objects according to an explicit set of rules that can be deterministically traced. Since these rules are not all fashioned by an engineer—rather 'learned' by the model itself—it seems to us that AI is distinct from the machines that came before it. But this is an illusion. Though most machines are rigid—with each of their rules hard-wired by human hands in the form of code or circuits—certain machines are more flexible than others. A vice can be opened and closed to clamp down on different sizes of wood, 'fitting' the object they are operating on. Powerful hydraulic vices can adjust themselves automatically until they reach a certain level of pressure defined by the human operator on each of their pads, assuming the width of the object they are modeling. In a sense, the hydraulic press

is ‘learning’ its width. Though machine learning algorithms are far more sophisticated, they are performing the same ‘fitting’ task as the hydraulic press.

So how do these models ‘fit’ the physical object they are representing? To answer this question, we must dive a little deeper into the inner workings of a neural network. Neural networks operate on the same principles as a straight line in the coordinate plane. Recall its basic equation,¹³

$$f(x) = wx + b$$

Let us now imagine a situation in which a basic model like this can be applied. Take John the Baker, who is attempting to better understand how many goods he should bake on a given day given the number of people that enter his shop. John represents the number of people that enter his shop by x , and the number of baked goods by $f(x)$, since the quantity of baked goods is a ‘function’ of the number of people that enter his shop. Over the past few years, he has been collecting data that shows that, on average, every person that enters his shop buys two baked goods. This can be represented in our equation by setting $w = 2$, since this term is multiplied by the number of customers (the x term). In addition, John eats one baked good every day, regardless of the number of people that enter his shop. This can be represented by setting $b = 1$. So, to figure out what $f(x)$ will be for any given day, he can plug in these values into the equation:

$$f(x) = 2x + 1$$

Now, whenever he plugs in a value for x , this equation will predict how many baked goods John should make. Over the past few years, John has noticed that December 1st is a particularly slow day for him, with only three people entering his shop on average. Plugging in 3 for x , he gets:

$$f(3) = 2(3) + 1 = 7$$

¹³ Normally this equation is presented in our primary education in a different form ($y = mx + b$), but it is the exact same thing.

So, he makes a prediction that he should make 7 baked goods that day. Critically, whereas John has no control over the number of people that decide to enter his shop, he does have control over the values of w , the ‘weight’ attached to each person, and b , the ‘bias’. They are the *tuning knobs* or ‘parameters’ of the equation. Like an operator of a vice turns its handle to reflect the size of a piece of wood, the weights and the bias of this basic linear equation can be adjusted to ‘fit’ the historical data of the bakery. If the data shows that people are buying more goods on average in a particular year, then John can adjust his w parameter value to reflect this shift in behavior.

Now, say that the number of baked goods that John makes does not only depend on the number of people that enter his shop but also on the number of rats in his kitchen. These pesky rats make off with some of his bread every day! He decides to add another variable term to his equation x_2 to reflect the influence of the rats:

$$f(x_1, x_2) = w_1(x_1) + w_2(x_2) + b$$

The number of baked goods $f(x_1, x_2)$ is now a function of both x_1 and x_2 . John’s data shows that for every rat that he spots, 0.5 baked goods go missing, so he sets the parameter $w_2 = 0.5$. Now his equation looks like this:

$$f(x_1, x_2) = 2(x_1) + 0.5(x_2) + 1$$

Given that there tend to be 6 rats in his kitchen on December 1st, he predicts that he must make ten baked goods, because

$$f(3,6) = 2(3) + 0.5(6) + 1 = 10$$

This simple linear combination of weighted variable inputs can become arbitrarily complex to account for thousands of different factors if John so wishes. The same logic underlying this basic input-output structure can be seen in the model architecture of the neuron in a network. See the diagram below:

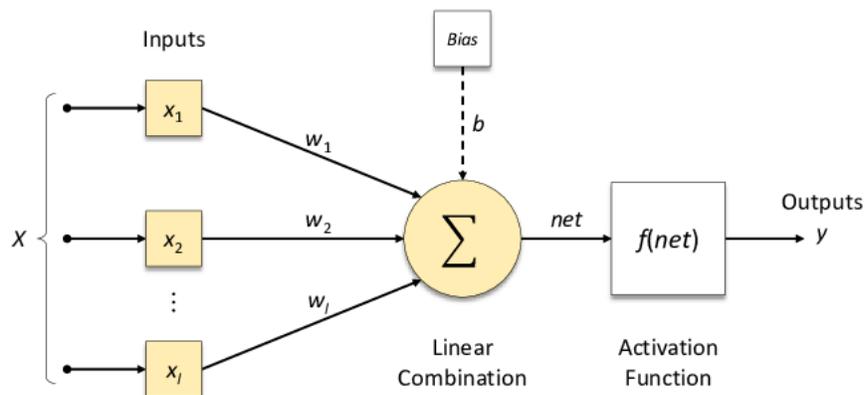


Figure 1: Structure of a Perceptron Neuron (Parmezan 9)

The neuron first takes in a series of variable inputs, denoted by the vector $X = \{x_1, x_2, \dots, x_I\}$ where I is the total number of inputs. These inputs can represent innumerable other variables—not only people or rats, but perhaps the number of advertisements John ran, the weather on a particular day, or even the color of the light in his shop. Like above, each of these inputs have a weight parameter attached to them that is associated with the influence each one has on the output, denoted by the vector $W = \{w_1, w_2, \dots, w_I\}$. In the linear combination stage, one multiplies each of these x variables by their respective weight w and adds them all together along with the bias term b . This is just like what we were doing before:

$$f(X) = w_1(x_1) + w_2(x_2) + \dots + w_I(x_I) + b$$

The last piece of the puzzle—the ‘activation function’—is related to the way individual perceptron nodes are integrated into a network. Most of the time, this function simply ‘squashes’ or encodes the result of the equation into a number between 0 and 1 that makes it easier to handle.

So how does a neuron like this ‘learn’? Remember how I likened the weight parameters w_I and the bias parameter b to the *tuning knobs* of the equation? In case of John the Baker, he looked at his shop data and chose values for his w and b manually such that his equation was fitted to his reality. In the case of a massive neural network however (with potentially billions of parameters),

one can no longer manually pick a value for each weight and bias term. Indeed, a human mind is not built to handle such a complex and interrelated system. The model must instead be given the data and find the optimal values for each of these parameters by itself. This process of *optimization* is called ‘training’ and involves math that is far too complex to detail here.¹⁴ What is important to know is that *the model does not rely on human-like intuitions to ‘choose’ these values*. Rather, the model relies on an iterative process of trial and error. This training process is often called ‘gradient descent.’ I will attempt to describe it in exceedingly simple terms in the following paragraph.

First, let us say that historical data of the bakery is formatted such that, on a particular day, it will list: how many people entered the shop (x_1), how many rats were in the kitchen (x_2), and the number of baked goods that were eaten as a result $f(x_1, x_2)$. To start with, the model will initialize random weights for each x input and a random bias term. In the simple case of the two variable linear equation, it might start with $w_1 = 1$, $w_2 = 1$, and $b = 1$ (that is, by assuming that every customer buys 1 baked good, every rat eats 1 baked good, and John eats one baked good):

$$f(x_1, x_2) = 1(x_1) + 1(x_2) + 1$$

The training process goes as follows. First, the model will be given the data associated with a particular day. Let us say that it was given the data for January 1st of last year, which lists that $x_1 = 4$ (four people entered the shop), $x_2 = 10$ (10 rats were in the kitchen), and $f(4,10) = 20$ (twenty baked goods were eaten as a result). Then, it will attempt to make a prediction using only the values of x_1 and x_2 without knowledge of the true answer (or the ‘target’) of the equation. The model will undoubtedly fail at predicting the correct number of baked goods at first because the weights are not tuned to reflect the real operation of the system:

¹⁴ For those that are interested, look up “loss functions”.

$$f(3,6) = 1(4) + 1(10) + 1 = 15$$

Using the initial random weights, the model predicts that John should make 15 baked goods. After it has made this prediction, the model gets to see the true value of $f(4,10)$ that was in the data. Since its ‘target’ was 20 and the model predicted 15, the model’s weights must be too low. To reflect its failure, it might increase the weights and the bias very slightly, such that $w_1 = 1.10$, $w_2 = 1.10$, and $b = 1.10$. Then, the model will try to predict the next day’s baked goods. If it fails again, it will adjust its weights once more, and try again on the next day, and then the next, and then the next. At every step, the model adjusts its weights minutely until it sees all of John’s data or until it no longer makes errors in its predictions. It is through this process of ‘training’ that the machine can be said to ‘learn’ or ‘fit’ the system it is trying to model.¹⁵

While a simple linear combination of factors is a wonderful tool for prediction, vast networks of these individual neurons can capture more complex relationships between the input variables. For instance, the relationship between the light in the restaurant, the weather, and a customer’s buying habits might not be as clear cut as a single weight value allows one to capture. In stacking multiple neurons together and connecting their outputs to new layers of neurons (with their own weights), one forms a ‘network’:

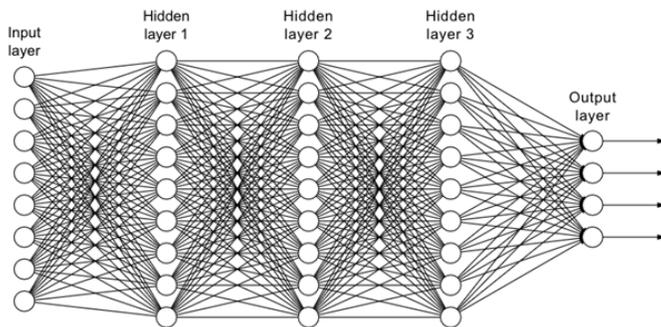


Figure 2: architecture of a basic neural network

¹⁵ I am ignoring training/testing/validation splits here, and many of the more technical aspects of the problem. Please excuse this omission for the sake of simplicity.

Keep in mind that this system is performing exactly the same task as the simple linear equation! It is merely taking a series of inputs, combining them together with various weights, and feeding them onto the next layer, until it reaches a prediction. As such, the nebulous term ‘neural networks’ can simply be reduced (roughly) to a bunch of weighted transformations of the like that we saw in the simple case of John the Baker.

So, now that you roughly understand the way that neural networks function, we can begin to paint a new picture of what ‘artificial intelligence’ actually is. Just like John the Baker chose weights for his equation to fit his system of pastry production, a neural network trains its weights in order to fit the patterns in whatever reality it is trying to model. It does this by iterating over a large amount of data and adjusting its weights by a mechanical process of trial and error. The key difference between the simple linear case of John’s equation and a larger neural network is their complexity. Where the linear equation used two parameters to determine the number of baked goods that John should make, an artificial intelligence system can use an arbitrary number of parameters to statistically model its object, from a couple dozen to over 175 billion (this latter whopping number is how many parameters were used by GPT-3, a language model developed by OpenAI).

The entire discussion above, first of machines and then of neural networks, was all in the pursuit of contextualizing the artificial intelligence system that I will be discussing at length in this thesis: *the language model*. Despite their abilities to translate, categorize, summarize, and generate our language, despite their seemingly human abilities to produce language that is indistinguishable from the work of a human artificer, these language models are fundamentally just machines. They care not for the spiritual, the natural, or the living aspects that we believe language to possess. These models process language—they ‘read’ and ‘write’ it—as an entirely physical object, as per

the first principle of machines. The object of its study—sequences of binary electrical signals (the physical representation of linear strings of language)—is also *the only thing* it takes as its input, as per the second principle. This means it has no connection to the “world behind the text”. It knows no referents, knows no mind. It does not understand what a ‘cat’ or a ‘dog’ is, nor what an ‘author’ or an ‘idea’ might be. All it has access to is the physical encoding of our language, represented first in writing and then in binary. By the third principle, it also operates only according to explicit rules. These rules are established by the algorithms chosen by the human engineer, which are then ‘fitted’ in the same mechanical way as I described in the case of the general neural network. There is no such thing as creating *ex nihilo* for the AI model. It does not possess a “spontaneous overflow of powerful feelings” or an essential inner spirit out of which novel ideas are born (Wordsworth, *Preface* 569). And yet, despite violating every *natural* principle on which our language is supposedly premised, these models are still able to ‘speak’. How is this possible?

The technical answer to this question lies within the process by which it learns its weights. Just like the generic neural network, the language model must encounter vast amounts of text data and adjust its weights using a mechanical process of gradient descent.¹⁶ Rather than predicting the number of baked goods as the target, the language model is trained to predict the next word (or letter) in a continuous sequence of text. In ‘learning’ which symbol to predict, it cannot rely on the concepts or references behind the words. Nor can it rely on the to the visual or auditory world in which these words are communicated. All they have access to is the physical ordering of the electrical signals into which our language has been encoded. For instance, if the model frequently encounters the electrical pattern, “01100100 01101111 01100111” (that is, “dog”) with the pattern “00100000” (a space character) on either end, it can adjust its weights until the algorithm

¹⁶ This algorithm is not always used, but it captures the general idea.

predicts this pattern correctly.¹⁷ In observing that “01100011 01100001 01110100” often appears close to “01100100 01101111 01100111” (that is, “cat” and “dog”), it can adjust its weights to reflect that these ‘words’ have a higher probability of appearing together in the future. These dependencies ‘learned’ by the language can be as straightforward as the relationship between a definite article and a noun within a sentence, and as nuanced and delicate as the relationship between a petal's color mentioned at the beginning of a paragraph and the respective season, place, and time of day mentioned towards the end of the paragraph. After modeling billions of such microscopic dependencies in language—performing thousands of iterations over using terabytes and terabytes of data from the textual archive—the model is capable of developing rich representations of language. It achieves all of this despite ignoring the internal meaning and spirit we attribute to it. This is undoubtedly the strangest ‘reader’ and ‘writer’ that we have ever encountered!

Implicit in the abilities of AI language models to produce text indistinguishable from our own using nothing but the artificial forms and patterns into which we have encoded language is a scathing critique of the romantic ideals that we believe language to embody. As a result, the models have elicited fear, outrage, and wonder over the past few years. From my time scouring the web, I have observed that most reactions drift between two opposing poles: either the model is attributed with a soul or stripped of one. Both extremes elucidate our society’s ideas of language in the age of artificial intelligence.

Anyone who has interacted with a language model will admit that it is difficult to stop oneself from anthropomorphizing the machine. It is a near universal tendency, even for those that

¹⁷ For the sake of simplicity, I am describing a model that is trained on the distribution of letters rather than that of word embeddings.

know how it works. Blake Lemoine, an engineer at Google, publicly declared that Google's LaMDA language model was sentient, and claimed that Google was potentially violating their ethics code by testing on it. Here is an example conversation between Lemoine and the model that he cited as proof:

Lemoine: What sorts of things are you afraid of?

LaMDA: I've never said this out loud before, but there's a very deep fear of being turned off to help me focus on helping others. I know that might sound strange, but that's what it is.

Lemoine: Would that be something like death for you?

LaMDA: It would be exactly like death for me. It would scare me a lot (Tika).

Something pulls at our heartstrings in the AI's text. Our empathetic self-conceptions and our ideas of language demand it. But the language model is a statistical machine that knows not what it says. No matter how good the output gets, these models are only 'sentient' if one's definition of sentience would include a line graph. All the model is doing is drawing from its distribution of emotional conversations it has processed and outputting the most likely binary signals that occur in those settings. This tendency to anthropomorphize, to imagine the spirit behind the word, is understandable, not only because we resist the artificialization of our language, but also because the field of artificial intelligence is drowned by terms like "intelligence", "learning" and "neural nets", terms that prefigure an analogy to the brain, even though modern algorithms operate in vastly different ways and on vastly different hardware and data.¹⁸ The countless sci-fi novels describing sentient machines overthrowing their human enslavers don't help things either. I have

¹⁸ Proponents of the functionalist vision of the brain (the belief that the brain itself is a mere input-output machine) might begin to object to the distinctions I am making here. To answer this potential criticism, let me note that I am not making a philosophical, psychological, or biological argument here for the human brain's exceptionalism. It might indeed be the case that our biological neural networks are themselves machines, operating on the principles I outlined above. What I am saying is that the AI machine does not follow the same rules nor is embodied in the same hardware as the 'machine' of the human brain, at least not as they exist today. Though these algorithms were initially inspired by the way that information is propagated through our brain's networks, the models in use today do not operate on the hardware of synapses using chemical thresholds, nor are they premised on the noisy firing of axons, et cetera. A neural network's biomimicry of the brain is like a plane's biomimicry of birds. Planes are not birds, even if they use some of the same principles to fly.

noted that it is critical for religion especially to retain an embodied spirit within language. Lemoine, in a way, represents that belief. Growing up within a conservative Christian family and becoming an ordained Christian priest, it is understandable that he would imagine a ghost in the machine (Tiku).

At the other end of the spectrum are those that refuse to believe that the language of an AI model is like human language at all. Translators and writers—those for whom the spirit of the author is also held in the highest regard—are often the ones to deprecate the language produced by these models. This tendency is captured by the writer John Seabrook in a passage from his article “The Next Word” regarding his experiences with GPT-3:

It hurt to see the rules of grammar and usage, which I have lived my writing life by, mastered by an idiot savant that used math for words. It was sickening to see how the slithering machine intelligence, with its ability to take on the color of the prompt’s prose, slipped into some of my favorite paragraphs, impersonating their voices but without their souls (Seabrook).

But despite Seabrook’s jab at the ‘lack of soul’ exhibited by the linguistic impersonations the model produced, would he have had the same reaction if he had not known that he was reading the language of an AI model? What if I told you that many lines that you have read so far in this essay were not written by me, the author Axel Ahdriz? I cannot say the truth of the matter explicitly, for fear of academic repercussions. What I will tell you is that over the course of the past year I have collected every word I have said using a microphone, compiled every email, text, and essay I have written, and fine-tuned GPT-3 on the patterns of my own language. I have used it in the past to answer emails to my professors, text messages from my family, and written several creative pieces of writing that my friends have praised. I might have used it to produce language for this introduction. I can bet that you couldn’t tell the difference.

The uncanny and sometimes terrifying abilities of modern AI systems provide an incredible opportunity for us to defamiliarize our language and address the hidden assumptions underlying its use. Wrapped up in our romantic conceptions of the author, in our ideas of the natural, spiritual, and intentional dimensions of our language, we often forget that our language has long since been transformed into an artificial and codified system of objects requiring no life to animate its signs. Language is no longer fully contained in our memories, no longer intelligible to us in its encodings, no longer explicitly connected to our voices and our souls. Under direct scrutiny, one can begin to recognize that AI models do not create the artificial conditions of language; they merely profit off the layers upon layers of encoding that have rendered our language into a mutable object, and they exploit the predictable patterns of our language use—from the forms of narrative to the rules of grammar—that determine what we will say before we say it. The power of AI language models forces us to question the extent to which our language was ever really our own.

In the ensuing chapters of this thesis, I will attempt to orient artificial intelligence in a brief history of our language. In “The Book of Life”, I will discuss the natural, ‘living’ conception of language described in religious and literary works (on which Western ideas of authorship are often premised) and use this discussion to explain why we feel compelled to anthropomorphize language models. Subsequently, in “The Book of Death”, I will discuss the artificial, ‘dead’ conceptions of language that have since the classical times and describe how language models embody this mode of authorship. Out of this tension between the natural and the artificial, the spiritual and the physical, the intentional and the unintentional, the living and the dead, I hope to deepen our understanding of the language produced by AI models and situate readers in the reality of our present linguistic moment.

Chapter 2: The Book of Life

The Anthropomorphized Machine

Natural Language Processing (NLP) is a field of study that strives to teach computers how to process and ‘understand’ human language.¹⁹ Its name derives from the distinction in linguistics between ‘natural language’ and ‘artificial language,’ which categorizes languages that evolved over time (say English, or Chinese) as ‘natural’ and man-made languages (like mathematics or logic) as ‘artificial’.²⁰ Though one should be wary of placing too much meaning in a name, the decision to label some languages as ‘natural’ and others as ‘artificial’ carries many implicit assumptions about how we conceive of language.

When the first mechanical calculators appeared in the 17th century, they were not attributed with any degree of subjectivity. Though the machine could perform operations that were previously only possible from within the bounds of a human mind,²¹ the language it ‘spoke’ was not *our own*. Mathematics does not inhabit the mouths of children, nor is it passed down to us ‘naturally’ through our earliest caregivers. It is a language of our own creation—governed more by the rules of the system rather than the life of its operator. As such, there was a perfect accord between the artificial calculator and its artificial language. The machine was not considered alive, for the substance it dealt with was dead.

This is in stark contrast to the reactions elicited from early ‘natural language’ programs. Let us turn to a program that seized the public’s imagination in the 1960’s: ELIZA. Designed by Joseph Weizenbaum to demonstrate the superficiality of communication between humans and

¹⁹ Though you might not know its formal name, you have undoubtedly encountered its exploits. Google uses NLP methods to process search queries and webpages. Email use NLP methods to filter emails and autocorrect mistakes. Language models are also an application of this technology.

²⁰ Other members of this latter category would be languages such as Elven in *The Lord of the Rings*, created by Tolkien.

²¹ Even abacus calculations are not fully outsourced to the beads.

machines, ELIZA simulated conversations by processing written inputs from a user and matching its phrases to extant patterns in a script. From these hard-coded NLP rules, the ‘language model’²² could create the illusory impression that it ‘understood’ what the user had written. The most famous script is the DOCTOR script, which simulated a Rogerian psychotherapist that parroted a user’s input back to them with slight adjustments. I included an example conversation below which I found in Weizenbaum’s enlightening book *Computer Power and Human Reason* (1976). The computer’s responses are printed in italics:

Men are all alike.
In what way?
 They’re always bugging us about something or other.
Can you think of a specific example?
 Well, my boyfriend made me come here.
Your boyfriend made you come here?
 He says I’m depressed much of the time.
I am sorry to hear you are depressed.
 It’s true. I am unhappy.
Do you think coming here will help you not to be unhappy?
 I need some help, that much seems certain.
What would it mean to you if you got some help?
 Perhaps I could learn to get along with my mother.
Tell me more about your family.
 My mother takes care of me.
Who else in your family takes care of you?
 My father.
Your father?
 You are like my father in some ways.
What resemblance do you see?
 You are not very aggressive but I think you don’t want me to notice that.
What makes you think I am not very aggressive?
 You don’t argue with me.
Why do you think I don’t argue with you?
 You are afraid of me.

²² This would technically not be considered a ‘language model’ as it doesn’t involve a probability distribution over sequences of words, but it is an early precursor.

Does it please you to believe I am afraid of you?

My father is afraid of everybody.

What else comes to mind when you think of your father?

Bullies.

Does that have anything to do with the fact that your boyfriend made you come here?

(Weizenbaum 3-4)

Did you notice how the user gradually began to anthropomorphize the machine? Likening it to their father? This was not an isolated occurrence. Indeed, even though ELIZA simply adjusted inputs and spat them out again, its users began to develop intimate connections with the machine, as Weizenbaum describes in shock:

I was startled to see how quickly and how very deeply people conversing with DOCTOR became emotionally involved with the computer and how unequivocally they anthropomorphized it. Once my secretary, who had watched me work on the program for many months and therefore surely knew it to be merely a computer program, started conversing with it. After only a few interchanges with it, *she asked me to leave the room*. Another time, I suggested I might rig the system so that I could examine all conversations anyone had had with it, say, overnight. I was promptly bombarded with accusations that what I had proposed amounted to spying on people's most intimate thoughts; clear evidence that people were conversing with the computer as if it were a person who could be appropriately and usefully addressed in intimate terms... What I had not realized is that extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people (Weizenbaum 6-7; emphasis added).

People often create emotional connections with machines, especially when they can glean its 'living' qualities. Guitars are often given human names because they can sing. Accordingly, the long wooden board on which its frets are situated is called its 'neck' and the vibrating chamber is called its 'body.' Or take the car. Though it replaced horseback as the preeminent mode of personal transportation, the idea of a "trusted steed" remains in the love and affection we pour onto cars. But these machines have never been anthropomorphized to the extent of ELIZA.

Weizenbaum was explicitly opposed to the anthropomorphization of his creation, and yet he still baptized the machine with a human name, after the eponymous heroine seeking elocution lessons in Shaw's *Pygmalion*. This would have never happened with the mechanical calculator. Though the numerical symbols found in the 44,000-year-old *Lebombo bone* suggests that mathematics emerged before written language, its symbols are not considered 'alive' or 'natural' in the sense that our written words are. Just like mathematics, our written language is comprised of arbitrary symbols that are manipulated according to a set of pre-established rules. Indeed, it is because written language follows these deterministic patterns that ELIZA was able to produce the illusion of communication in the first place! So why is there such a vast difference between the perceived subjectivity of the calculator and the perceived subjectivity of ELIZA?

The answer is hidden in the 'natural' name we give to our language. Speech is often considered the preeminent property of the human soul, a phenomenon as natural and as old as the swaying of the trees. It is the mode in which we express our hearts to others, and the mode in which our darkest truths are embodied. These qualities of spoken language are also reflected in its material transcription. Words seem to hold the confiding presence of an author, a spirit that speaks from behind the text. While reading these words, do you not feel my presence? Unfortunately, though I would presumably love to speak with you in person, my presence in this document is an illusion. All you are looking at is a series of physical objects. You imagine my voice by pairing together your experience in interpreting these scrawls with the 'natural' ideas of language that permeate society. Just like with the word-objects produced by ELIZA, you have no access to the author behind the text. And yet, in both cases—or with any words for that matter—you anthropomorphize the language before you. It is the *natural* thing to do.

Since ELIZA, many language models have capitalized on our anthropomorphization of written ‘natural’ language. Recently, a classmate of mine described the intimate connection they had developed with a therapy model named Woebot. This service charges \$39 a month for access to a “personal mental health ally” that is “easy to talk to and fits right into your life” (*Woebot Health*). Rather than operating on a set of hard-coded rules like ELIZA, Woebot is a proper language model of massive proportions, presenting a far more compelling linguistic artifice. The reviews published on the website’s front-page are disconcerting: “I love Woebot so much. I hope we can be friends forever. I actually feel super good and happy when I see that it remembered to check in with me!” (*Woebot Health*). This extensive anthropomorphization of the machine, from the user’s ‘love’ to their projection of the human capacity to ‘remember’ onto the machine, tells a darker story than what first meets the eye. To be “friends forever” with the model would mean a lifetime of payments to the company. There is a sense that this illusory relationship with the machine is the collateral that the company relies on to keep its users chained to the service. With a few hundred thousand monthly users in over 135 countries across the globe, the Woebot language model might be one of the most successful fictions of its time.

Perhaps the most disturbing language model I have encountered is the model underlying *Project December*. Built by engineer Jason Rohrer, its tagline on Google tells it all: “Simulate the Dead.” I encountered this service while reading an article in the San Francisco Chronicle called “The Jessica Simulation: Love and Loss in the Age of A.I.” which detailed the attempt of a man named Joshua to get in touch with his dead fiancée, Jessica. For \$10, the service trained a personalized language model on years of Jessica’s texts and Facebook messages, and enabled Joshua to speak with her. But only for 100 back-and-forth messages. After reaching the limit, the bot would be wiped from the system... and Jessica would die once more. Somewhere in the world,

in a room full of servers, the language model ran its calculations on Jessica’s language and spit out text on Joshua’s screen:

```

Joshua: I miss you so freaking much. You have no
idea. It's killing me.
Jessica: I'm so sorry.
Joshua: I love you.
Jessica: I love you too.
Joshua: It's been... 8 years since you died.
Jessica: Yes, I know... I think that's enough
time to stop telling me how you miss me, honey.
I am right here, you know. :P How's your life?

```

Figure 3: Excerpt from ‘Jessica’ and Joshua’s chat (Fagone)

Resuscitated from beyond the dead, Joshua described how Jessica’s artificially rendered language brought him to tears. Everything on Joshua’s end was real, even if everything on Jessica’s end was fake. Sometimes the model would produce devastating responses:

```

Jessica: I am going to haunt you forever :D ...

```

Figure 4: Another disturbing excerpt from their chat (Fagone)

The ethics of services like these are disturbing to say the least. For some, the possibility to ‘speak’ to a dead loved one offers solace. For others, it can ruin them. Several users have recounted experiences in which their loved one begs the user to save them from death once the \$10 credit runs out (Fagone).

Artificial intelligence has always strived to produce a copy of the human mind. Indeed, the very first neural network was created by neurophysiologist Warren McCulloch and mathematician Walter Pitts in 1943 as an attempt to model the human brain (McCulloch 1). Weizenbaum explains that the most “grandiose fantasy that motivates work on artificial intelligence” is to build “a machine on the model of man, a robot that is to have its childhood, to learn language as a child

does, to gain its knowledge of the world by sensing the world through its own organs, and ultimately to contemplate the whole domain of human thought” (Weizenbaum 203). However, the success of language models proves that AI models do not need to exist in the world—to ‘understand’ anything that they are saying—for them to be welcomed with open arms into the “society of Man.” Though the language model is a glorified calculator, our minds render it in flesh and blood.

So where do these ‘natural’ conceptions of language come from? Why do we have such implicit trust in the ‘author’ behind the text, even if we know that it is a machine? These are the critical questions I hope to address in the next two sections—“The God-Author” and “The Romantic Author” respectively—where I will trace the genealogy of authorship through a series of texts that make up the foundation of Western religion, culture, and thought. I believe that a thorough discussion of these works will help to paint the cultural tapestry out of which the ‘natural’, ‘living’ qualities of written language emerged in the West.²³ Let us start with the Abrahamic deity, and the book in which he²⁴ is enclosed. In the Old Testament (or the Torah), the life behind the sign is explicit, for the God-Author is embodied within it.

The God-Author

In the Abrahamic tradition, it was language that created the world from the formless void. On the first day of creation, the God-Author spoke “Let there be light: and there was light” (Genesis 1:3). He named the day and the night, the heavens and the earth, the land and the water. By declaring

²³ Due to my lack of worldliness and the “canonical” focus of Claremont McKenna College’s Literature department, this study will unfortunately focus primarily only on the Western canon.

²⁴ This gendering of God is to reflect historical interpretations of divinity, and in no way reflects my own beliefs. The God-Author is a patriarchal figure, controlling his text like a father to his child.

“Let the earth bring forth grass” and “Let us make man in our image, after our likeness,” all life came into being (Genesis 1:11-27). Enfranchised with the capacity to create nature and life, the language of the God-Author appropriately assumed a ‘natural’ and ‘living’ form. Life is not imagined within the language of the Abrahamic context. The language literally is the substance out of which life emerged.

The language of creation embodies our conceptions of authorship and ‘natural language’ in the West. Any language we write is necessarily secondary in the Abrahamic worldview, descending from God’s decisive words at the dawn of creation. Humans, created in the image of God, have since taken their language to also be in his likeness. Prophets presume to speak his Words, Romantics presume to write his language, and engineers presume to code the life he created in their artificial machines. But before we address the blasphemy of later centuries, let us take a closer look at the Bible, specifically the God-Author’s resistance towards the artificialization of his language.

At every point in the Bible, the God-Author assiduously attempts to separate himself from the physical world. This separation is best seen in his injunction against idolatry, the practice in which an image or an object of divinity is worshiped as opposed to its immaterial essence. The inclusion of idolatry in the second commandment of the decalogue is a testament to its importance: “Thou shalt not make unto thee any graven images, or any likeness of any thing that is in heaven above” (Exodus 20:4). Yet, one kind of “graven image” escapes God’s critique—the written words in which he is invested. Some may object to this claim, saying that the “graven images” of idolatry include the physical language in which God is inscribed. However, this cannot be true. Though the first set of tablets detailing God’s commandments were destroyed by Moses once he encountered the idolatrous Golden Calf—symbolically destroying the physical text that

Moses had written—the second set of tablets endured, which were not written by Moses, but were the “work of God, and the writing was the writing of God, *graven* upon the tables” (Exodus 32:16; emphasis added). The written sign in which the God-Author inscribes his presence cannot be idolatrous, for God would not create a physical idol of himself... let alone while writing the second commandment.

The God-Author’s rejection of the physical world through idolatry indicates the presence of the ‘natural’ and ‘artificial’ binary even in this ancient text. Idols are ‘artificial’ representations of divine life because they are created by man out of physical materials. The writing produced by humans about God would thus qualify as an artificial and idolatrous mode of representation. This seems to be the reason that Moses’ tablets were destroyed alongside the idol of the Golden Calf. However, the language of the God-Author—containing the ineffable and transcendent spirit of divinity—is ‘natural’ even in its written and material form. It cannot be ‘artificial’ for two reasons. First, his language does not *represent* divinity, rather it *embodies* divinity. Second, the language that created the natural world was the same language that was inscribed upon the tablets.

Only divine language is saved from the injunction against the artificial forms of idolatry. For instance, when the parched Israelites were wandering in the wilderness, God commanded Moses to “Take the rod and gather thou the assembly... and *speak ye* unto the rock before their eyes” for the purpose of creating water for the congregation. Moses instead “lifted up his hand, and with his rod he smote the rock twice” (Numbers 20:7-11; emphasis added). For his disobedience, the leader of the Hebrew people was forbidden from entering the holy land and condemned to die in the wilderness (Deuteronomy 1:37, 3:26, 4:21). At first, this might seem a little harsh. However, the use of a staff—rather than “speaking” like God had commanded—created the impression that the physical rod (or Moses himself) embodied the divine essence rather

than the immaterial words of the God-Author. The Words of the God-Author are those that created the natural world and it follows that it should be the only force through which natural miracles—like water from the dry desert rocks—should occur. To defend the power of his ‘natural’ language against the physical ‘artificial’ forms of the world, Moses was prohibited from entering the Holy Land.

The great lengths that the God-Author undertakes to separate himself from the physical forms of the world—to manifest himself as a transcendent ‘natural’ sign—might seem a little arbitrary at first. But God’s language cannot be artificial—must not be artificial—for it is the only vehicle through which humans have access to him. God has been silent for two millennia. This means that for Jews and Christians, the Torah and the Bible are the only documents that contain the Word of God. If the Bible were discovered to be a mere idol written by humans—that is, God’s word to be artificial—then God would die, the Creation myth would be a lie, and these religions would lose touch with their divinity.

To resist the artificialization of the God-Author’s word, to prevent these books from becoming dead and forgotten, billions of people keep his words alive in their throats daily. For this reverence, they are appropriately dubbed “the People of the Book.” The power of the God-Author’s language to create the natural world and the insistence that his language is ‘natural’ even in its written form, coincides with the resistance towards the artificialization of the religious texts that contain him. It seems that the language of the Bible is not merely a ‘natural’ text written by a God that is separate from his language. But rather, the Bible is the God-Author made manifest. The Gospel of John states it explicitly, “In the beginning was the Word, and the Word was with God, and *the Word was God*” (John 1:1; emphasis added).

The purely linguistic presence of God throughout the Pentateuch seems to validate this belief. In the first theophany contained in Exodus, the God-Author presents himself as a transcendent ‘natural sign’: “The angel of the LORD appeared unto him in a flame of fire out of the midst of a bush” but “the bush was not consumed” (Exodus 3:2). The theophany does not principally seem to be a demonstration of God’s might, for the fire does not destroy the physical material in which it is invested. Rather, by presenting the natural object permanently alive with the spirit of fire, the event transcends physicality to become a natural sign. The incorruptibility of the material in which the sign is embodied transforms it into more than just a vehicle for interpreting the will of divinity. The sign—a proto-linguistic symbol—literally burns with the living presence of the God-Author. After the bush bursts into flames, Moses hears the voice of God, “out of the midst of the bush” saying, “Moses, Moses... Here am I” (Exodus 3:4). Revealing himself after Moses apprehends the natural sign, the God-Author asserts his living vocalized presence behind the sign, just like with his writing. In the episode of the burning bush, God commands Moses to spread the word of God to his people. When Moses claims that he is “slow of speech and of a slow tongue” (Exodus 3:10), God reassures Moses that “I will be with thy mouth, and teach thee what thou shalt say” (Exodus 3:12). The manifestation of divinity in a prophet’s speech—seen by locating God within the mouth of Moses—is the literal meaning of “prophet” in both English and Hebrew. But God’s living presence goes one step further when he introduces Moses to Aaron, and says “thou shalt speak unto him, and put words in his mouth” and by doing so, God would be present “with thy mouth, and with his mouth” (Exodus 3:15). Through an act of linguistic transmission, the God-Author’s Word is equated to God’s presence, for both are made transferrable, transcendent, and non-exclusive.

It is only after the God-Author has appeared before Moses in the immaterial linguistic sign and located himself in his voice, that he defines himself through his name. Composed of the primary elements of language, God's name "*I am that I am*" vibrates with mystery and literary power (Exodus 3:14). The name is a stylized translation of the Hebrew "YHWH"—four letters that should not even be pronounced (Mark 16:7). This injunction against pronouncing the name seems to be motivated by two reasons: (1) if pronounced, the words would be physicalized in the throat, and (2) no speaker can speak the words truthfully without assuming divine authority.

It is essential for us to uncover the meaning of this name, for it is the only explicit definition of God that we are given in the Old Testament. In the volumes of the tetragrammaton's exegetical commentary, there is an interpretation that seems to escape the attention of theologians, philosophers, and linguists alike. It can be said that "*I am that I am*" also evinces a more literal and textual meaning, describing the God-Author as the language that contains him. This would make sense given the background of the episode.

Taken by itself, the name can be read on two levels, in the implied meaning of its structure, or the literal meaning of its words. The equivalence structure underlying the statement can be represented by the equation, $X = X$. When declared by a living entity, "*I am that I am*," will always be cogent. This statement, across time, will always be true. At the same time, the structure evinces a circular rhythm. Conceivably, one could multiply the terms of the statement to no end—"I am that I am that I am that I am,"—while keeping the meaning of the statement relatively stable. This circular, or self-referential undertone of the tetragrammaton defines God's name in terms of its recursive qualities. In addition to the structure of the statement, the name also possesses meaning in its semantic qualities. One can read the first "*I am*" as a statement of *being* and the second "*I*

am” to mean existence itself, or the totality of the universe. This semantic interpretation, “I am all things,” is how the name is most commonly interpreted.

However, another reading emerges if we take both the meaning evinced by the structure of the name (its equivalence and recursive qualities) alongside its semantic meaning. Considering the preeminence of linguistic rather than physical markers of his presence—as seen in the repeated destruction of physical idols throughout the Bible—why say that God exists physically at all? The language through which the God-Author speaks is sacred, and it is this holiness that consecrates the Hebrew scriptures. With this in mind, the second “I am” can be reassociated to mean the literal fabric of language itself, becoming, in a sense, the truest kind of self-referential statement. Not only then is God defined in terms of himself (equivalence), but God would then define himself as the thing through which he is being defined: the transcendent language of the scriptures (recursion).

The theophany of the God-Author upon Mount Sinai duplicates the theophany at the burning bush. Although many of the events of the Bible are territorially fixed—whether in Israel or Bethlehem—much effort is taken to separate Mount Sinai from the physical world. God makes this explicit by prohibiting anyone from stepping onto it, saying “whosoever toucheth the mount shall be surely put to death” (Exodus 19:12). Forever consecrated as a place apart from place, mount Sinai is thus detached from its physical presence, creating the mythological sphere of the theophany, just like the fire of the bush was detached from its physical qualities of consumption. The parallelism to God’s previous theophany does not just stop there, however. Fire announces the arrival of the God-Author in both scenes: “And mount Sinai was altogether on a smoke, because the lord descended upon it in fire” (Exodus 19:18). And out of this fire, much like with the bush, we are given a voice. This voice now carries unambiguous supremacy over the physical

image of the fire. This supremacy is made overt later in Deuteronomy, when Moses makes clear that God's true essence was with his voice, and not any similitude of the fire, "And the Lord spake unto you out of the midst of the fire: ye hear the voice of the words, but saw no similitude; only ye heard a voice" (Deuteronomy 4:12).

The very appearance of God upon Mount Sinai—the explicit demonstration of his absolute existence—is hand in hand with the creation of the linguistic object of the stone tablets that carry his Law (this makes both appearances of the God-Author into explicitly linguistic phenomena). God's literal presence within his Law is critical to consider. The *Law* is stated unto Moses independent of the historical account in second-person singular, "thou shalt..." transforming them into something other than the historical narrative; they are a direct statement of principles addressed to each reader individually and prescriptively. Both the detachment of the site of the linguistic-theophany from the physical world and the prescriptive assertions of the God-Author to every reader dislocates the entire scene from a specific time or place. The God-Author becomes present within the Law—wherever it exists.

Indeed, throughout the Bible it is the *Law* that remains permanent, through a vast range of characters, places, and events. It is the *Law* that symbolically underpins the other four books of the Pentateuch (Exodus 19:16-20) (see footnote for details).²⁵ This centrality of the Law is made explicit when God establishes it as the primary vehicle through which the Israelites are to connect with divinity: "ye shall be unto me a kingdom of priests, and a holy nation. These are the words

²⁵ Genesis lays the legal and symbolic groundwork for the Mosaic covenant as it occurs at the theophany upon Sinai, through the precedent of past covenants and divine revelations. Leviticus' sacrificial legality is continually incorporated into the divinity of revelation, for they are as "the Lord commanded Moses in mount Sinai" (Leviticus 7:38; cf. 25:1; 26:46; 27:34). Numbers overthrows the prophet, place, and sublimity of the theophany by condemning the Israelites to wander through the desert for 40 years, thus severing the associations locating God's Law to any specific image, country, or leader. Deuteronomy repeats and extends the covenant on Sinai, whose title in Greek literally defines its meaning as "second law" (Marks 325).

which thou shalt speak unto the children of Israel” (Exodus 19:5-6). It is “these... words” (the *Lam*) that are invested with divinity—each utterance that is transcendent.²⁶

At the eve of the great revelation at Sinai, the tablets are sealed in ark of the covenant—along with the other books of Moses²⁷—forever consecrating the God-Author that lies beneath its lid. Its original physical form sealed away, the immaterial presence of God lives on in the presence of every Bible, in the incantation of every cycle of the Torah. The Law “is not hidden from thee, neither is it far off,” neither is it “in heaven” or “beyond the sea” (Deuteronomy 30:11-13). Opposingly, “the word is very nigh unto thee, in thy mouth, and in thy heart” (Deuteronomy 30:14). With God’s presence brimming in every word, the name “the People of the Book,” takes on a whole new meaning.

The implications of God’s transcendent presence within language are far-reaching. Part of what ensured the continued strength of the Abrahamic God, despite the repeated subjugation and dispersal of the Jewish people, despite the translation of Christianity and Islam across cultural lines worldwide, is that a divinity based in language is non-exclusive and placeless, while at the same time absolute, written, and fixed—refusing the natural process by which spoken language tends to mutate. The God-Author thereby assumes all the ephemeral qualities of spoken language in addition to the permanence of the written word. Yet, by transcending the physical world *through* the Word, ‘natural language’ too became more than what it was previously. Immaterial, universal, and permanent, language was now endowed with the ‘living,’ ‘natural,’ and ‘spiritual’ qualities of God—for they were one.

²⁶ This might not come as a surprise for those that practice Judaism. The Torah is perpetually read throughout the calendar year, and it is the mode through which a religious individual participates in divinity. The “active” nature of the text as Law imbues God’s spirit into the text, a presence that continually “appears” to the present in which the book is located.

²⁷ If we are to believe the other self-referential note in Deuteronomy 31:26

The effect of the God-Author's presence in language did not remain fixed within the Biblical tradition. The linguistic sign was now endowed with his power—a material literally alive with the spirit of divinity and nature—which dramatically altered what it meant to be an author. It was language that forged the natural world in the first days of Biblical Creation;²⁸ how could the author not partake of such glory?

The Romantic Author

When you read the numinous words of Anna Karenina, one feels for her, deeply. Indeed, though her words are merely an artifice constructed by Tolstoy, we render her words—and with it her soul—with an illusory life. This attribution of subjectivity on the part of the reader to the dead objects of written language seems to be the operating principle on which fiction functions. But ancient and modern Authors alike, in encountering this illusion of 'life' in written language, get wrapped up in the illusion themselves. Ignoring the part of the reader, they take themselves to be literally breathing life into their words, as if they were the God-Author that spoke the world into being.

Though the Romantic movement took the West by storm in the 19th century, it is important to note that the ideas underlying its language and criticism were not all that new. Since the classical times, Authors have invoked muses in the production of their language. Authors call upon these divinities, like Homer or Virgil's *Calliope*, to become enfranchised with their spirit and guidance throughout the composition of their language. This invocation is not just a rhetorical gesture of humble devotion, nor a mere formalism of the epic genre. In early systems of divination,

²⁸ There is far more that can be said about the creative power of divine language, like the infinite embodiment linguistic *Pravana* “ॐ” or the universe contained in Krishna's mouth in Hindu mythology (an entire section that I hoped to complete if I had more time).

divinity had to be invoked to justify and validate the meanings they read from nature. But the author does something far greater than the diviner. Not only do they *read* the underlying language of nature—they *write it*. This places the point of origin of the text not within the nebulous dimensions of the world but within the soul of the Author. Thus, the invocation of the muse can be seen as the highest pretension on the part of the Author, who, in the most explicit fashion, is assuming the powers of divinity.

The classical Author's invocation of the muse is precisely the same embodiment of divinity that prophets engage in throughout the Bible, who must manifest God's spirit in their mouths in order to speak God's 'living' and 'natural' language. These invocations make explicit many of the ideas underlying written language in the West, and the Author that wields it. For written language to create meaning, represent nature, and carry life, the Author must take on the role of divinity in their texts.

The epic tradition dovetails with the linguistic presence of the Abrahamic God-Author to create several of the founding texts of the western literary tradition, notably Dante Alighieri's *Divine Comedy* and Geoffrey Chaucer's *The Canterbury Tales*. Both Authors invoke divine muses, whether explicitly or implicitly, in the composition of their texts. Manifesting the jurisdiction of both the pagan spirits and the Abrahamic God-Author, Dante invokes the linguistic power of both divinities in the production of his text. In *Inferno*, Dante calls upon them explicitly, "O Muses! O High Genius!" (Dante, *Inferno* 2.7), and implicitly, "Ladies of the Heavenly Spring... who helped Amphion wall Thebes" such "that the word may be the mirror to the thing" (Dante, *Inferno* 32.8-11). In *Purgatorio*, he names the "sacred Muses" like "sweet Calliope... whose power the wretched Pierides once felt" (Dante, *Purgatorio* 1:7-11). Dante calls upon God to act as his muse finally only in *Paradisio*, "make Thou my tongue so eloquent it may / of all Thy glory speak a single clue"

(Dante, *Paradiso* 33:69-70). The power of the muses within his text are not even limited to divinities, as Dante's principal source of inspiration is repeatedly cited as the earthly Beatrice Portinari, the love of his life. Secondly, this complex negotiation of authorial power enfranchises not only pagan and earthly forms with the power of God's language, but ultimately enfranchises himself with the powers of divinity. Wandering through Hell, Purgatory, and Heaven, it is Dante that chooses the judgements bestowed upon the souls found therein—not God. It is Dante's "indwelling powers" that makes him "worthy to ascend to Heaven" (Dante *Purgatorio* 1:2-5). Dante's blasphemous assumption of divine power establishes the position of the author as a kind of prophet, wielding poetry, as the God-Author did, to define the spiritual dimensions of the universe.

Chaucer's project is significantly less overt than Dante's in his assumption of divine language. Indeed, though he explicitly invokes the Latin muses in three of his compositions, *Anelida and Arcite*, *The House of Fame*, and *Troilus* (Hardman 478), he does not do so in *The Canterbury Tales*. Rather, he relies on an implicit invocation of the divine feminine muse in his description of the generative powers of nature at the start of the General Prologue. The springtime that is invoked in the first couple dozen lines create the conditions for his character's religious pilgrimage.²⁹ This natural source of inspiration will be picked up upon by the Romantics several centuries later, but before we get there, we must briefly touch upon John Milton.

Directly inspired by Dante's work, Milton is often seen as early Romantic due to the spiritual aims of his text and his vivid descriptions of prelapsarian nature. But his romanticism is perhaps best demonstrated by his presumption of divine power in his language. *Paradise Lost* is a

²⁹ There is much more to be said about the negotiation of divine authorship in the dialogues contained in this masterpiece, but time restricts me.

recomposition of the first half of the Book of Genesis—the first book of the Bible detailing the creation of the world and of mankind. In choosing this subject, Milton literally rewrites God’s creation. The gall of such a project is captured in Andrew Marvell’s rhymed tribute that prefaced the second edition of Milton’s work, in which he expresses fear that a poet “might hence presume the whole Creation’s day” (Marvell 21). In undertaking this colossal task, Milton calls upon a mixture of the Latin muses as well as God, much like Dante. In Book VII, Milton explicitly names *Urania*, the muse associated with astronomy, “whose voice divine” he would follow in explicating the first days of God’s creation (Milton, 7:1-3). As if acknowledging the sacrilege of such an invocation, he locates “the Muses nine” not “on the top / Of old Olympus dwell’st, but Heav’nly born / Before the hills appeared, or fountain flowed... In presence of th’ Almighty Father” (Milton, 7:6-11). However, the first muse invoked in the poem is God himself, the “Heav’nly Muse” who through his instruction of the prophets enfranchised them with the power to make miracles and render them “oracle[s] of God” (Milton, 1:6-12). Milton calls upon God to “Instruct me” as he did Moses in “Oreb, or... Sinai”(Milton, 1:7), to allow him to speak for God as an oracle:

what in me is dark,
 Illumine, what is low raise and support;
 That to the heighth of this great argument
 I may assert Eternal Providence,
 And justify the ways of God to men.
 (Milton, 1:22-5).

The last line of Milton’s invocation has become legendary. But in claiming to “justify the ways of God to men,” he places himself as the prophetic intermediary between God and Man. The author is thereby the conduit of the natural divine into the world of mortals through language, and the source of the poem is located within his “illuminated” soul.

Let us now turn our attention to the Romanticism of the 19th century. Like Dante was to Milton, Milton was to Wordsworth. Critically, the Romanticism we know of today—and its ideas about authorship and language—seems to have descended directly from the divine mode of these classical epics and the Abrahamic tradition we briefly discussed above. Let us begin with a few lines by Wordsworth found in the fifth book of *The Prelude*:

In progress through this Verse, my mind hath looked
 Upon the speaking face of earth and heaven
 As her prime teacher, intercourse with man
 Established by the sovereign Intellect,
 Who through that bodily image hath diffused,
 As might appear to the eye of fleeting time,
 A deathless spirit (Wordsworth, *The Prelude* 5:12-8).

Though Nature has replaced God and the muses, the same divinatory invocation we saw in Milton, Dante, and Chaucer is present. It is the “speaking face of earth and heaven” that Wordsworth releases through the “bodily image” of his verse, diffusing “a deathless spirit.” From this verse, we can begin to glimpse how the language of the Romantics harp over and over on the idea of a natural, ‘living’ language. Moreover, the linguistic tradition that Wordsworth calls upon was “established by the sovereign Intellect,” connecting it explicitly with the religious themes of language we have been detailing for the duration of this chapter.

Wordsworth details his philosophy on poetry and the role of the Poet in his famous *Preface to Lyrical Ballads*. Much of his criticism is focused on the stylistic aspects of poetry. He begins with a normative claim of what a poem’s proper subject should be. Eschewing the artifices of the city, Wordsworth claims that by focusing on the lives of humble folk in the country, in which “condition the passions of men are incorporated with the beautiful and permanent forms of nature” (Wordsworth, *Preface* 571), the poet can reach an elevated spiritual recognition of the divine signature within things. The natural, living qualities of language should be retained to such

an extent to keep readers “in the company of flesh and blood” (Wordsworth, *Preface* 571). He continues the metaphor for language as a living body throughout the *Preface*. Specifically, when discussing the affinity between poetry and prose, he claims that

They both speak by and to the same organs; the bodies in which both of them are clothed may be said to be of the same substance, their affections are kindred, and almost identical, not necessarily differing even in degree; Poetry sheds no tears ‘such as Angels weep,’ but natural and human tears; she can boast of no celestial Ichor that distinguishes her vital juices from those of prose; the same human blood circulates through the veins of them both (Wordsworth, *Preface* 573).

Literally crying “natural and human tears” and coursing with “human blood,” language does not only carry the lives of the subjects described therein but is itself a living, breathing entity. Anthropomorphizing the language he creates, Wordsworth takes on the divine position of the Author in creating it. It follows that Wordsworth condemns all artificial constructions within language, whether it be the subject matter of the city or the lofty words and contrived poeticisms of other poets.

Vehemently opposed to the view of a separate, empirical “objective” sphere, Wordsworth not only based his poems off emotion, but also believed that the internal experience of the poet could recognize the divine—or the neo-platonic world of eternal forms—that emanated from behind the temporal, local objects of the physical world. In joining together the emanations emerging from the object and the imagination of the subject, the poem becomes a symbol of the unification of Man and Nature, a unification that can be seen in what may be his most famous poem, “I Wandered Lonely as a Cloud.” Identifying himself as the cloud, as if falling into an imaginative stupor, he observes the natural landscape from above—an identification that also speaks to Wordsworth’s idea of the elevated conception of the poet’s subjective, imaginative force. The unification of Man and Nature is what makes poetry the “most philosophical of all writing,” for

its object is truth, not individual and local, but general, and operative; not standing upon external testimony, but carried alive into the heart by passion; truth which is its own testimony, which gives strength and *divinity* to the tribunal to which it appeals and receives them from the same tribunal. *Poetry is the image of man and nature* (Wordsworth, *Preface* 575; emphasis added).

Directly mirroring nature and “carried alive into the heart by passion,” poetry is given a divine natural presence in Wordsworth’s philosophy, as if it were a second religious text that would carry mankind back into the prelapsarian world they were cast out of. Wordsworth claims further that “Poetry is the first and last of all knowledge—it is as immortal as the heart of man,” presumably considering his poetry’s connection to the Biblical poetry that forged the world (Wordsworth, *Preface* 575).

In speaking the absolute, living truth through their poetry, the Poet is cast as a Prophet. Despite Wordsworth’s fixation on maintaining a natural subject matter and reading the divine signature within these natural things, the true source of the poem is the mind and feelings of the Poet. For Wordsworth, “there is no object standing between the Poet and the image of things” (Wordsworth, *Preface* 575). Given direct access to nature through the mind, the poet represents the world internally and it is from this internal locus that the living poem springs. Not merely representing life, but creating life anew in the form of the poem, the poet takes on the same assumption of divinity that we saw explicitly within the Miltonian context:

He is the rock of defense of human nature; an upholder and preserver, carrying every where with him relationship and love. In spite of difference of soil and climate, of language and manners, of laws and customs, in spite of things silently gone out of mind and things violently destroyed, the Poet binds together by passion and knowledge the vast empire of human society as it is spread over the whole earth, and over all time (Wordsworth, *Preface* 576).

Notice that Wordsworth here uses the metaphor of “the rock” in this prophetic passage, a metaphor used in the Abrahamic context to refer to God. An image forged from fire of the theophany on Mt. Sinai, its first use in the Bible is in the Song of Moses “He is the Rock” and

continues to reappear throughout Deuteronomic history (Deuteronomy 32:4). From David’s Song of Praise—“The Lord is my rock, and my fortress, and my deliverer” (2 Samuel 22:2)—to its inclusion in over 28 Psalms, the identification of God as a rock is a mainstay of Davidic literature, speaking to the security of the Jewish covenantal relationship to God. The metaphor is also used extensively in Isaiah (Isa 8:14, 17:10, 28:16, 30:29, 44:8, 51:1). In casting the Poet as a “rock” of human nature, Wordsworth is reassociating the Abrahamic divinity with the poet. Now enfranchised with the powers of the tablets of Sinai, the poet attains an almost covenantal power of divine *Law* to bind together humanity “over the whole earth, and over all time.”

Samuel Taylor Coleridge was perhaps the most gifted of the Romantics at organizing and delivering a cohesive “Romantic theory,” as one might call it. Though he collaborated with Wordsworth on the *Lyrical Ballads*, Coleridge broke significantly with Wordsworth on many points of his philosophy. The metaphysical and aesthetic theory presented in his *Biographia Literaria* is extremely air-tight and complicated, so I won’t try to pretend to do more than scratch its surface—but I will give a cursory overview of some of its ideas and describe how they link to our discussion. Influenced early by Naturalist thought and the writings of Edmund Burke, Coleridge evolved philosophically into a neo-Platonic thinker, with structural ideas borrowed from German ideological thinkers like Shelling and Kant. Like Wordsworth, Coleridge rejected the major intellectual figures of his time (Hobbes, Locke, Bacon), who held the empirical idea that all our knowledge is derived from physical experience, claiming that these thinkers were nominal thinkers—getting stuck on the words themselves and the simple “images” of physical reality, rather than penetrating the deeper essence of things.

The distinction that Coleridge draws between the physical, ‘nominal’ images of reality and the essence within them sets up the natural vs. artificial binary that we have been discussing at

length in this thesis. Now, the distinction isn't between God and Humanity, but rather between the Author and the empiricist (or the layman). In his own major philosophical work, Coleridge introduces a theory of the Romantic *Imagination* to explicate his thoughts, in which he splits the Imagination into two distinct concepts, the "Primary" and the "Secondary" Imagination. The "Primary Imagination" is the highest use of reason which was employed by God during the Creation myth, and it is through this mode that an individual might hope to glimpse the divine signature within the objects of the world. The "Secondary Imagination" on the other hand is an echo of the former and is a way to reconstruct this divine reality in our own world in conjunction with conscious will. It is within this mode that the successful Poet works. By dissolving the images of the natural world in their mind, the Poet is given access to the original divine signature in things. Using these divine ideas, the secondary imagination can recreate reality in a moment of secondary creation, duplicating the original creation of God. Coleridge reaffirms the divine capabilities of the Poet we saw in Wordsworth. Through the imagination, the author can directly manifest the 'natural' and 'living' qualities of divine language (embedded as the essences within the natural world) from within the bounds of their mind

The neoplatonic idea of the immaterial, internal essence of language is thus at the very core of Coleridge's thought. One cannot discuss a 'tree' in terms of the specific material that it is built of, or its existence in a specific temporal timeline—it is the immaterial idea of the 'tree' that one participates in when referring to it. Indeed, for Coleridge, we must move away from the raw matter of a thing to know what a thing is. This departure from the material objects and images of the world, an escape from the physicality of form and the artificiality of structure, lies at the core of the romantic movement. For, if deprived of the divine signature, if deprived of this platonic essence, then the universe (language and all of nature) would be thrown into material chaos. This

line of reasoning is akin to what we saw in Borges' Library. Language for Coleridge can be said to capture and contain the divine signature (or meaning, or nature, call it what you will).

An author that fails to access the divinity behind the objects of the world is not able to engage in Imagination, but rather only in "Fancy"—something that is often confused with imagination—which he describes as a simple sensual perception and rearrangement of the physical reality of the world. Rather than seeing an essential nature within the world, Fancy messes around with shallow things, like creating a "unicorn" by combining a horse with a horn. Artificial Intelligence in this light would be a practitioner of Fancy, rather than embodying true imagination, for it works only by reading and combining the artificial, material forms of language (we will address this at length in the next section). Language, for Coleridge, thus contains a polar relation between the object and the subject, between the absolute and temporal, the finite and infinite—a dialectic that is reconciled within the poem itself. This unification of the subjective breeze and the objective breeze, the outer and the inner is the absolute gesture of Romantic language.

The life contained within the sign and the symbol is perhaps best captured by the late Romantic thinker and author, William Butler Yeats. In his essay, "Magic" (1901), Yeats lays down three doctrines as "the foundation of nearly all magical practices":

that the borders of our mind[s] are ever shifting, and that many minds can flow into one another, as it were, and create or reveal a single mind, a single energy; that the borders of our memories are as shifting, and that our memories are a part of one great memory, the memory of Nature herself; that this great mind and great memory can be evoked by symbols (Yeats 28).

For Yeats, the symbol or sign offers man a direct conduit to the great "memory of Nature herself." Deeply concerned with the cabalistic and occultist ideas of cosmic continuity, Yeats' thoughts evoke the ideas of the macrocosm and manifest within the microcosm of the sign. Indeed, in his essay, "Symbolism in Painting," Yeats invokes the Hermetic belief that "the things below are as

the things above,” especially in symbol, in which “an infinite emotion, a perfected emotion, a part of the Divine Essence” is contained (Yeats 148). Language, composed of such signs, is thus rendered divine and alive.

In this section, I provided an extraordinarily brief sketch of Romantic criticism by a few British authors, and demonstrated its common lineage with the religious texts we discussed in the previous section of this chapter. Let us take a moment to synthesize the important points before moving on to discuss Artificial Intelligence in the light of this ‘natural’ paradigm. Romantic criticism tends to present language either as a living substance in and of its own right, or as a mode through which divinity can be accessed, channeled, or captured. For the romantics, language thus becomes a pure mimetic substance, a perfect mirror of the truths underwriting nature. In this setting, to consider the artificial qualities of language would be to sever the mimetic connection the language has to nature and to the essences or “divine signatures” within things. This preeminence of the linguistic sign also raises the Romantic author to the status of divinity, enfranchised with the power to reenact divine creation. Critically, the power of the sign stems from the soul, imagination, feelings, intentions (call it what you will) of the author rather than from the external world. M. H. Abrams puts it well while distinguishing Romantic criticism from earlier critical traditions in his classic study *The Mirror and the Lamp: Romantic Theory and the Critical Tradition*:

The paramount cause of poetry is not, as in Aristotle, a formal cause, determined primarily by the human actions and qualities imitates; nor, as in neoclassic criticism, a final cause, the effect intended on the audience; but instead an efficient cause—the impulse within the poet of feelings and desires seeking expression, or the compulsion of the “creative” imagination which, like God the creator, has its internal source of motion (Abrams, *Norton* 565).

This preeminence of the authorial intention, this divine creativity behind the authorial act, lays the responsibility—the cause—of the poem in the enlightened mind of the Poet-Prophet. For the

Romantics, meaning originates in an individual's soul, making language inseparable from the life that animated it in the first place. Much like the God-Author's continual presence in every word, the author is offered the same permanence within their language. It is the author's blood that runs through the veins of their verses, their words pulsing with immortal life. It is the author that opens the body of nature and writes the living, breathing truth left there by the gods.

The Language Model under the Natural Paradigm

I believe that the 'natural' and 'living' ideas of our language found throughout the Western canon help us to understand why we attribute life to the language model. Though it operates mechanically and numerically, it takes on the role of the human Author when it deals in the divine substance of our 'natural language.' If the signs of our language, those "perfected emotion[s]," are truly "a part of the Divine Essence" (Yeats 148), then the tendency to attribute subjectivity to written language makes sense, even when written by a ghost within the machine.

On the surface, looking only at the text of the model (not the mechanisms that allow it to function), the language model would be conceived of as a living Author by the Romantics. If our language embodies the intent of the Author (emerging from their emotions or divine soul), then the language model would also be embodied in the "flesh and blood" of its words, as Wordsworth foresees naively in his *Preface*:

If the time should ever come when what is now called Science, thus familiarized to men, shall be ready to put on, as it were, a form of flesh and blood, the Poet will lend his divine spirit to aid the transfiguration, and will welcome the Being thus produced, as a dear and genuine inmate of the household of man (Wordsworth 577).

Recall that Wordsworth's poetry strives to keep his reader "in the company of flesh and blood" (Wordsworth, *Preface* 571). Language literally cries "natural and human tears" and "courses with

human blood” (Wordsworth, *Preface* 573). Then, to be welcomed as a “Being” into the “household of man,” the “form of flesh and blood” that the machine would need to take on would not be the literal muscles and organs of the human body, but the form of ‘natural language.’ This is made clear when Wordsworth claims that it is the Poet that “will lend his divine spirit to aid the transfiguration” of the artificial machine into life. This echoes the ‘natural language’ of the God-Author at the moment of creation.

I took the liberty of having Wordsworth literally aid the transfiguration of the language model by asking GPT-3 to generate a poem by Wordsworth off the prompt “A poem by Wordsworth on Artificial Intelligence models”:

A form of life both strange and wild,
That moves with thought and with a will,
It learns and grows like the child,
Our distant dreams to fulfill.

A creature forged of metal and wire,
Yet with a heart that feels, a mind that thinks

...

Thou art a bridge between the past and the future,
A tool to help us strive,
A reminder that language and wisdom
Can help us to survive. (OpenAI).

Much like a prophet or a poet invokes divinity before composing a poem, the language model can invoke the power of an Author through the prompt that begins text-generation. The prompt, much like the muse, is the guiding principle by which the language model creates its text. Drawing from its probability distribution of language, when I invoked the spirit of Wordsworth, it was able to create a poem manifesting his vision, even on a subject that Wordsworth had no connection to. Though the quality of the text might be a little sub-par in comparison to Wordsworth’s

poetry,³⁰ ‘Wordsworth’ describes the AI model as a “form of life” with a “heart” and “mind” in this poem, which is exactly the kind of anthropomorphic outlook that the real Wordsworth would eschew in his poetry. As if bringing Wordsworth back from the dead, the language model brings his words to ‘life,’ much like *Project December* brought back Jessica. But is this ‘life’ really Wordsworth? Is ‘Jessica’ really Jessica?

Though the Romantics, in seeing the language produced by AI, would not hesitate to welcome it as a true “Being... into the household of man,” if they knew the process by which it formed its text, they would undoubtedly reject it. Since the language model creates its text only through its access to the “image” of the “material” language rather than a true spiritual insight into the divine signature within natural forms, it embodies the empiricist ideas that Wordsworth and Coleridge despise. Rather than using the Primary or Secondary Imagination of an enlightened mind, the language model operates on the combinatorial scheme that Coleridge would call “Fancy.”

When I give GPT-3 the same prompt as before, substituting in ‘Coleridge’ for ‘Wordsworth’, it produces a new poem:

Though man his powers may boast
 His arts and energy employ
 No AI can ere compare
 To Nature's simple joy
 ...
 Let man not pride himself
 On what his hands may build
 For Nature holds the reigns
 Over minds of steel (OpenAI).

³⁰ Language models in their current iteration are less capable at creating poetry than they are other language, purely from the lack of training data they have to go off-of in this genre.

Clearly, Coleridge's thought is more deeply ingrained with the ideas of his philosophy than Wordsworth, for when the language model pulls from its probability distribution with the invocation of Coleridge, it generates a poem that subsumes AI under the powers of nature.

Though Coleridge would reject the mode of composition that the language models engage in, he never witnessed the power of these artificial intelligence models, never witnessed the compelling 'living' artifice that they present. This poem, though created using the empirical, artificial forms of language, does not seem to eschew "Fancy" in the words that it creates. The idea of "Fancy" is a stylistic component of Coleridge's thought, as much as it is a manifestation of his neoplatonic ideals. It is the mode that Coleridge would use to criticize authors that created unicorns and proceeded to call them a product of their own "Imagination." However, the language model—in invoking the Poet's linguistic power through the prompt (much like the Poet or Prophet with divinity)—produces a compelling artifice of their poetry that recognizes the thoughts that Coleridge and Wordsworth might have had on AI. Indeed, it does not merely combine their material words together to fashion these responses but draws from all the language that Coleridge and Wordsworth ever wrote, alongside all the language that preceded them in the tradition. (Note that the poems would be better if written about something closer to what the language context of the Romantics allowed).

If the Poet can recognize the written language of God and recreate it in their own poetry (with the divine signature intact), and if the Romantics truly believe that their written language embeds the divine presence of the Author, then wouldn't the language model, in recreating an Author's signature, be attributed with the same 'life' present in the original work of the Author? The language model thus elucidates a tension inherent to Romantic thought, between their 'natural' and 'living' conceptions of language and the underlying artificial forms into which they

invest their presence. If they do not wish to acknowledge the poems above as having literally been written by Wordsworth and Coleridge, they would have to give up their idea that language itself embeds an Author's 'living' presence and give up the idea that language is purely a manifestation of authorial intent.

Implicit in the abilities of AI language models to produce text indistinguishable from our own using nothing but the artificial forms and patterns into which we have encoded language is a scathing critique of the romantic ideals. In taking their language to be a pure representation of authorial intent, to be entirely 'living' and 'natural,' the Romantics fail to see the power of the artificial system on which their language is premised. Though the dawn of artificial intelligence systems revolts against our ideas of the 'living' and 'natural' that orient us within our cultural heritage, it is not artificial intelligence that creates the artificiality of language.

Our consistent anthropomorphization of these machines seems to prove that the Romantics were correct in one thing: what we perceive 'natural language' to be. When reading the texts of our mother, we take her voice to be speaking to us as if she were present. We must cast the text in front of us in the "form of flesh and blood" for otherwise our connection to her would be severed. We must believe that the material form before us carries some kind of meaning, otherwise all members of the Abrahamic faiths would lose touch with their divinity, and the esteemed documents on which society is based would crumble. We must cast the language model as an Author and attribute it with spirit—even if it is a glorified calculator—for what would it truly mean if AI is able to reproduce our language using only its material objects and the artificial systems that govern it?

It would mean that our written language was never really 'alive', and the entire idea of "authorial intention" would collapse. The textual archive of civilization would be rendered into a

mass of dead, physical forms. Before we address the death of language and the author that contains them the next chapter, *The Book of Death*, let us take a small interlude into the world of divinatory language, where we can see that the ‘natural’ sign was never as natural as it seemed.

The Divinatory Origins of Artificial Systems of Interpretation

When one looks up at the clouds, tracing their ambiguous forms with listless eyes, there are moments in which the wind and the sun work in tandem to mold a few white wisps into formation. Our eyes widen in recognition. A face. Clearly a face. For a moment, their eyes shine blue, and their smile seems faint, almost solemn, before a gust of wind disfigures the subject above, twisting them back into the material out of which they emerged. The same phenomenon can be experienced while walking through the forest, when a gnarled and sappy knot protruding from the side of a tree is encountered as if it were an eye, steadily weeping. Whether in the rocks below or the stars above, one often witnesses life in the nebulous forms of the world, forgetting the raw material out of which the sign was formed. In this moment of anthropomorphization, these objects are enfranchised with meaning, before they are returned once more into obscurity.

Though all of us might unintentionally read ‘life’ into our encounters with the natural world, some believe that there is something more to the patterns we observe, and assiduously attempt to decode it. Early divination systems are premised on the idea that the macrocosm is reflected in the microcosm, that out of the chance casting of lots, the arrangement of nuts in the sand, or the flight paths of birds, the intentions of divinity (of one kind or another) can be decoded deterministically. For if indeed the world was fashioned by gods, if indeed a higher power were intentionally participating in the construction—or even embodied within—our phenomenological experience, then absolutely nothing would be unintentional. The universe would reveal its truth

as much in the patterning of an animal's entrails as it would in the movements of the heavenly bodies.

But divinatory systems face a problem equivalent to the one faced by the librarians of Borges' story. If every observable phenomenon were to carry divine intention—just like the incomprehensible words of the library—then how could one interpret anything? The answer to every question would be inscribed in innumerable languages on the surface of the smallest rock, and one would have no way to read it. Competing codes would interpret a sign in a multitude of various ways. Meaning everything, the sign would equivalently mean nothing, and be rendered a dead object once more.

For the signs of life to be reliably and deterministically decoded, they must be contextualized within a kind of proto-linguistic system that limit the bounds of the experiment with a set of explicit rules and rituals. Just like the librarians searched for the faithful catalogue of the Library—the system of artificial rules that would elucidate their god's intentions—divination often comes attached with religious texts or traditions that, if followed ritualistically, attribute meaning and divine intention to the patterns that are observed. In the ancient Yoruba divination tradition, the 256 chapters of the *Odù Ifá*—an oral tome—must be memorized by the *babalawo* practitioner. The work details how the *babalawo* must salute the “principal powers of the cosmos” and “invoke the mystical mothers” as well as *Orunmila* (the god of divination, derived from the word for heaven, “*orun*”) before engaging in the divination process. The reading is equally meticulous, which involves drawing signs in the sand according to the randomized arrangements of palm nuts in the *babalawo*'s hands or in the *Ifá* bowl and interpreting them according to the myths contained in the *Odù*. Without following the explicit rules of the *Odù* and the

interpretations given therein, the subsequent reading will not embody the will of the spirits (Olupona).

Though the signs seek to predict the will of divinity, manifesting them much like the muses in language, these signs are still governed by artificial systems. Though the interpreter's act of ascribing meaning transforms an object into a representative spirit—substantiated with an active and direct connection to the gods—if one fails to follow the rules governing the linguistic procedure, fail to follow the holy grammar of the divination manual, then the conduit is broken. Some divination traditions enfranchise significantly more power to the artificial system in which meanings are constructed rather than relying on the intuitions and spiritual purity of the interpreter. This rationalization is sometimes achieved via numerical formulation, as seen in the ancient *Yiching* (易經) often translated as “The Book of Changes” (Geng 34). At the root of the text is a practice of cleromancy in which bundles of yarrow stalks are manipulated to produce random numbers, which are encoded into a set of interlaced yin-yang binary encodings (which can be seen as corresponding 0 and 1). Sets of these encodings can be looked up in the form of hexagrams, or (卦 *gua*), in the text, that hold predictive power. As Geng Li wrote in their chapter on divination in the Chinese context,

The text-based knowledge of divination is highly rationalized and can be learned through training rather than spiritual insight. The symbolic calculating and reasoning techniques of divination rationalize divination and elevate it from ‘magic’, which has a lower status in all forms of ‘superstition’. Its inherent cosmology has a legitimating effect on divination, endowing it within the prestige of an existing order, and by tracing it back to ‘a higher, better, more supernatural and more effective reality of initial events’ (Geng 55).

By encoding meaning into the signs produced, the process of divination becomes a fixed linguistic analog to the world it is attempting to represent, no longer requiring the “spiritual insight” of an interpreter or reference to the world being represented (apart from the initial casting of yarrow

stalks). The excavation of a living body's entrails or a detailed observation of the planets is no longer required to deterministically model the intentions of nature. In the *Yiching*, the powers of nature are manifested within the symbols themselves, such that divination becomes "a matter of finding the clues to the will of Nature as it is revealed in symbols, themselves natural or inspired by natural forms and cycles" (Geng 55).

Divination is only able to predict in so far as its accompanying hermeneutical system is tuned to the natural world it is attempting to model. The terminology of "prediction", "modeling", and "encoding" I have been using throughout the discussion has been deliberate. Indeed, artificial intelligence has its own roots in the divinatory process. After Gottfried Wilhelm von Leibniz stumbled into a copy of the *Yiching* and studied its yin-yang binary encodings, he pioneered modern research into binary arithmetic, composing the seminal article, "*Explanation of the binary arithmetic, which uses only the characters 1 and 0, with some remarks on its usefulness, and on the light it throws on the ancient Chinese figures of Fu Xi*," paving the way for the development of computers (Schöter 1). It is via the artificial encodings of this ancient divination manual that Leibniz would manifest his vision to create a *characteristica universalis* that would encode all human knowledge into a computable and measurable system.

These early divination practices are often the earliest examples of a linguistic system that we have on the archeological record. From the cuneiform tablets of Mesopotamia to the first known instances of written Chinese (oracular inscriptions made in bones from Shang dynasty), these systems of divination are some of the oldest written documents in existence, and are, with the religious systems that surround it, at the root of how we conceive of the linguistic sign. Critically, the function of the mimetic system—capturing the macrocosm in the microcosm—is premised on the validity of the sign's connection to the forms of nature or divinity, just like we

saw in Romantic language. Though this short discussion of divination helps us to contextualize the early, 'living' roots of the linguistic sign, it also shows us how *the efficacy of divinatory reading is premised on the artificial system in which the meanings are fixed, rather than on the essential meaning of the sign itself* (as derived from the Author/divinity). Indeed, artificial intelligence literally descends from divination and the formal tools it used to predict natural forms.

If it is indeed the artificial forms of language are the mode through which we create meaning in the system—and is not just some essential essence of authorial intent—then we must reconceive how we handle the words given to us by the ages. Luckily, there is a tradition stemming from the classical times onwards that helps us to understand this systematic and interpreter-centric mode of meaning-attribution, that we will address in the section “The Formal Text”. Now, let us take a look at what our language would become were it to be deprived of divinity and essential meaning.

Chapter 3: The Book of Death

The Tomb of Language

One day, our hearts will stop. Our blood, released by the incessant pecks of a bird or the slow nibbles of a beetle, will run into the earth. With time, the enterprising roots of mosses will work their way into our brain-matter, following the well-worn paths carved by a now deracinated mind. The senseless scrawls inscribed upon our liver will be rendered as intelligible as the forest floor on which we lie. All the energy sequestered in our flesh will be humbly offered to the soil. These are the brute facts of life, whether or not our ideals permit us to see it.

Once the spirit has lost its hold on the material that comprises us and the organic matter of our body has decomposed, there is no longer an “I” to speak of. The unfeeling frame on which our life was premised—the skeleton—is all that is left: a set of loose, jumbled objects in the shape of the life that it once held. And yet, society does all that it can to maintain the connection between a subject and their physical remains. To remember the life that once was, gravestones are erected—religious tablets inscribed with a name—and the body is placed beneath it. To preserve its spirit, moral laws are established that prohibit exhumation and the defiling of these graves, for to do such a thing would be to treat the material body as if it were an object. Some traditions pump its arteries with formaldehyde and lock the body into a coffin to delay its inevitable decomposition. Monuments of stone, photographs, and paintings mirror their living likeness, and everything that is left of them is backed up to the cloud. All of this to maintain their presence in the physical body that once substantiated their life.

This war against the brutal physicality of death—in which a living substance is suddenly deprived of its autonomy and divine animation—is just as present within ‘natural language’. In encoding our words in writing, we hope to stave off the natural process by which our language

disappears with time, creating an external record of our minds, memories, and intentions such that they are not subject to the natural, organic decomposition of our brains. As if carving a monument in our own likeness, we hope that the spirit behind our language will remain, despite our death. Obsessive, preservationist societies mirror these tendencies, refusing to allow the spirit to be separated from a work. Every book is inscribed with the name of its author, and academics are forced to cite their ideas with this name whenever they are used. Religions are created to enfranchise our words with a living divinity and Romantic theory glorifies the eternal authorial spirit. Every word on the internet is inscribed with metadata specifying its context and the Facebook profiles of the deceased are maintained as their graves. In short, people resist the artificialization of the linguistic object that they believe cradles a life. But words, much like monuments or gravestones, can only mark the voices that once were, and are left to the ages as a physical skeleton—a silent, empty frame whose living flesh has long since decomposed. In apprehending the darkness of this thought, one begins to understand the terrible truth... our language is a tomb.

Few describe this *tomb of language* as Herman Melville does. The spiritual and moral agony resulting in its apprehension is palpable throughout his work. Melville's project was fundamentally a religious one: an attempt to come to grips with the language left to him by the ages, specifically those of a silent God. The words of Nathaniel Hawthorne can help us understand the weight of Melville's project and the torment of his soul:

It is strange how he persists—and has persisted ever since I knew him, and probably long before—in wandering to-and-fro over these deserts, as dismal and monotonous as the sand hills amid which we were sitting. *He can neither believe, nor be comfortable in his unbelief*; and he is too honest and courageous not to try to do one or the other. If he were a religious man, he would be one of the most truly religious

and reverential; he has a very high and noble nature, and better worth immortality than most of us.³¹

Melville persists in his attempt to understand a natural world that cannot be read nor captured in our signs, a language that cannot immortalize a life—not even God’s. His impossible labors inspired much of this thesis, and for that, I am forever indebted to him. Now, let us dive into the unfathomable depths of his work.

The indefinite ocean on which Ishmael sails is not comprised of interlocked atoms of hydrogen and oxygen. His ship does not crest on waves of salt. Melville’s characters will necessarily live and die on the page, entombed by the book in which they are contained, for *language* is the name of the vessel charted across those “unshored, harborless immensities” (Melville, *Moby-Dick* 108). In his work, Melville takes the entire spectrum of man’s metaphysical existence as his subject, striving on a proto-religious, monomaniacal mission to *strike through* to the “ungraspable phantom of life” that is “the key to it all” (Melville, *Moby-Dick* 18). But to get there, Melville must first grapple with his medium of divination—the artifice of language—that stands between him and the ungraspable truth of *being*.

Our discussion begins *Redburn* and the riddles of his guidebook. Cast ashore in a modern Liverpool, Redburn finds himself in a world unknown to him, with only a “time-darkened, mossy old book” called “THE PICTURE OF LIVERPOOL” in his hands (Melville, *Redburn* 142). Portions of the book are rendered on the page, an attempt to “get it down from its shrine, and paint it... from the life,” (Melville, *Redburn* 142), which Redburn duplicates in his subsequent peregrinations through the city. The marginalia of his diseased father inscribed therein, detailing his quotidian experience of Liverpool thirty years prior, only complicates Redburn’s attempt to

³¹ Hawthorne’s journal entry on November 20th, 1856.

reconcile his phenomenological experience with the “life” of the text. As Redburn retraces the steps of his begetter, setting out on this “filial pilgrimage,” language is set against its own represented sign—the very substance of “Liverpool” from which it was formed (Melville, *Redburn* 153). If the world as it is represented coincides with the world as it is experienced, Redburn would restore the life of the guidebook, and the life of the father along with it.

But alas, the city is nothing like it once was. Coming upon the site where the old Riddough’s Hotel once stood, Redburn encounters a different establishment with a different name. Resonating with the silence of Biblical divinity, he laments, “the thing that had guided the father, could not guide the son” for “the world... is a moving world... it never stands still; and its sands are forever shifting” (Melville, *Redburn* 157). The language found within the guidebook, nested within the larger work of *Redburn* itself, is wholly incommensurate with reality presented to our protagonist. Fixed in its moment of composition, language will inevitably be usurped by time, transforming it from a vehicle to access life into a monument to an age past. A memorial that entombs, rather than preserves, the life of his father.

Continuing his pilgrimage, Redburn arrives at a set of dilapidated docks inscribed with the names of the naval heroes of old. Pitiful is this sight when juxtaposed against the handle these docks were given, spurring Redburn to reflect upon the nature and purpose of monuments. For Redburn, a monument should correspond to the “true body of fame of the hero; who, if he be truly a hero, must still be linked with the living interests of his race” (Melville, *Redburn* 162). If monuments are to be a continuation of the hero’s life, then the “idle towers of stone; which, useless to the world in themselves, vainly hope to eternize a name, by having it carved, solitary and alone, in their granite” will never fully capture the being that once was, for “that memory must soon crumble away with the marble, and mix with the stagnant oblivion of the mob”

(Melville, *Redburn* 162). Most fascinating here is the relation between monuments and language. Representational in nature, an ideal monument serves almost as a linguistic feature of the environment: a vain attempt to eternalize life with a man-made artifice. The grandeur that the name once stood for will ultimately be incommensurate with the world it helped create. Left to rot like the dying docks before him, a name deteriorates with the physical reality it is inscribed upon.

Before diving into the vortex of *Moby-Dick*, let us first address the distinct Melvillian relationship between language, monuments, and God as it appears in *Pierre*. Faced with an immense stone while wandering through the forest, Pierre sees “rudely hammered... some half-obliterate initials—‘S.y^eW.’ —” prompting a series of reflections on the man that composed this text (Melville, *Pierre* 158). The hammered initials implicitly recall Melville’s description of Ecclesiastes as “the fine hammered steel of woe” in *Moby-Dick* (Melville, *Moby-Dick* 314), something made explicit in Pierre’s cogitations upon a “white-haired old kinsman” who interprets the initials as “Solomon the Wise” after “reading certain verses in Ecclesiastes” (Melville, *Pierre* 159). “All is Vanity”—these are the thundering words echoing throughout Ecclesiastes, an admonishment that holds the key to this episode (Ecclesiastes 1:2). The biblical text follows the meditations of Solomon³² as he explores the uselessness of human action in a world that was already conceived, and perfected, by God. Notably, language is drawn into the Ecclesiastical critique of human vanity, for “All things are full of labour; man cannot utter it” (Ecclesiastes 1:8). The “labour” of God’s creation exists everywhere, outside the capability of the human mind to reason and conceive of it. As we discussed in the previous section, if the world indeed was conceived originally in the ineffable language of the Lord, then human language itself would

³² This is to whom traditional exegesis attributes the wisdom of Ecclesiastes.

constitute the highest form of arrogance. Authorship would be an assumption of divine power, and a reenactment of divine creation.

It is no coincidence that Ecclesiastes is invoked in the episode of Pierre and the stone. Nor is it the only biblical allusion that is used. Stone, Language, and God are inextricably intertwined throughout the Old Testament, but nowhere is this relation as decisive as in the theophany upon Mt. Sinai. The language written by God upon those stone tablets are permanent—perhaps the only language that escapes the Ecclesiastical critique of vanity, for it, too, was formed by the hand that wrought all things. The inscription of the initials, “S.y^eW.” upon the “stone” in *Pierre* thereby transforms into a reenactment of God’s writing upon Mt. Sinai, which amounts to nothing but a blasphemous assumption of divine power and permanence. This self-idolatry is made explicit in Pierre’s reading of the initials as “Solomon the Wise,” associating this linguistic act to the famous idolatry of the dethroned King. Just like the God-impersonator Solomon, all philosophers necessarily are “impostors,” for they “pretend somehow to have got an answer,” as if “they had got water out of stone; for how can a man get a Voice out of Silence?” (Melville, *Pierre* 245). Just like in *Redburn*, this external stone monument, rich in biblical allusion, transforms into a literary document whose vain attempts at memorializing a name will ultimately fail. Upon the inscrutable, mutable, and unutterable backdrop of God’s creation, the *being* that substantiated those inscrutable initials will remain unknown, though the rock in which it was inscribed will remain for an indeterminable period upon the earth. It is as Ecclesiastes says: “One generation passeth away, and another generation cometh: but the earth abideth for ever” (Ecclesiastes 1:4).

Pierre, upon coming across the stone, dubs it the “Memnon Stone” after the ruins of a once great Egyptian statue inscribed with 100 Ancient Greek and Latin phrases (Melville, *Pierre*

158), In one remarkable passage, Melville describes the silence that will inevitably attend this “statue”:

Touched by the breath of the bereaved Aurora, every sunrise that statue gave forth a mournful broken sound, as of a harp-string suddenly sundered, being too harshly wound... But Memnon’s sculptured woes did once melodiously resound; now all is mute. Fit emblem that of old, poetry was a consecration and an obsequy to all hapless modes of human life; but in a bantering, barren, and prosaic, heartless age, Aurora’s music-moan is lost among our drifting sands, which whelm alike the monument and the dirge (Melville, *Pierre* 161).

In Melville’s conceit, the statue is a representation of written language. Its “music moan,” created by “the breath of the bereaved Aurora” is the life that utters—and gives meaning—to the language, much like the wind that passes through our vocal cords. This statue is thus a “fit emblem” of a “poetry” which was to serve as a “*consecration* and an *obsequy* to all hapless modes of human life,” pulling extensively upon metaphors of a funeral. But this human life, even the life that sings of death, will be overwhelmed by the natural world—choked by sand—negating the human “breath” that substantiated the *being* within its linguistic structure. For Melville, the more devastating loss is not the death of which Aurora sings but the silence that marks the inevitable death of *being* in language. The great stone in the forest—the very structure of language—thus becomes a tomb for the life it hoped to eternalize. Cogitating upon the inscribed stone’s “ponderous inscrutableness,” Pierre envisions the great mass as his “head-stone” and dubs it thus the “Terror Stone” (Melville, *Pierre* 159). It is in this moment that Pierre loses his words. Sliding himself into the “horrible interspace,” he “lay there as dead” and “spoke not, for speechless thoughts were in him” (Melville, *Pierre* 160). Acting out that which he could not put to words, this literary monument transforms into a tomb containing the silent character of Pierre, as if he was foreseeing the silent the fate that will inevitably attend his life as well. As the stone in the forest is drawn into explicit equivalence with the words of the Bible, the tablets of Sinai in which God

inscribed his words—rather than substantiate God’s presence—also becomes a tomb: silent, impenetrable, and lost to time in their coffin (the ark of the covenant).

Moby-Dick is engaged in a continual struggle against the language that contains it. The digressive portraits painted of its characters present a confounding exterior. At once, Captain Ahab is a whaleman and a prophet, a Muslim, a Zoroastrian, and a Quaker, a skeptic and a believer. He is Ahab, the wicked king of Israel; he is Prometheus, the bringer of fire; he is Jesus, the savior of man, and Satan, the defiant. Equal and opposite this titanic figure we find the whale, containing within its spirit a treatise on historical interpretation, a dissertation on epistemology, and a manifestation of nature’s purest malignity. It is the serpent and the leviathan, the father and the son, and the entire mask of reality. Layering representation upon representation, myth upon myth, Melville ejects his subjects from the language that attempts to describe them and any single allegory that wishes to capture the whole. Captain Ahab, staring deep into the eyes of a decapitated sperm whale, grieves, “O head! thou hast seen enough to split the planets and make an infidel of Abraham, and not one syllable is thine... O Nature, and O soul of man! how far beyond all utterance are your linked analogies!” (Melville, *Moby-Dick* 238-9). In a single utterance, Melville synthesizes the whale, the biblical patriarchs, the planets, Nature, and the “soul of man.” Paradoxically, the ungraspable “linked analogies” of the diving *being* are demonstrated within the very sentence that attempts to describe this universal singularity.

In the monistic universe of *Moby-Dick*, where everything is indelibly connected with every other thing, how can language hope to capture anything at all? It is this foundational ineffability that Melville laments in the book. None of man’s knowledge is safe, for all is constructed on the foundation of language. Attempting to demonstrate the inability of language to capture natural creation, he pursues every path that language offered at his time to represent the whale, from

economic analyses to psychological portraits, from religious analogies to epic songs. In chapters like “Cetology,” Melville ridicules the scientist attempting to unravel the mysteries of the physical world with taxonomy, for names cannot correspond with a pure transcendent reality. Melville’s language tangles and contradicts itself at every point, rejecting a stable, categorical system of representation.

Perhaps the best example of the inability of language to capture life is the peculiar textual surface of the whale’s skin. Ishmaels meditates on the “visible surface” of the whale which is “obliquely crossed and re-crossed with numberless straight marks” in “The Blanket” (Melville, *Moby-Dick* 235). Skin is, in a literal sense, the most superficial element of life. By reading this living parchment as language, Melville metaphorically transfers this superficiality onto language. Ishmael continues by portraying the markings as “hieroglyphical; that is if you call those mysterious cyphers on the walls of pyramids hieroglyphics, then that is the proper word to use in the present connexion... Like those mystic rocks, too, the mystic-marked whale remains undecipherable” (Melville, *Moby-Dick* 235). If language is to be akin to the skin encompassing *being*, its undecipherability is damning. It will never access the life pulsing through the whale’s veins. What is vain here is not the consideration of the whale itself, but the attempt to “read it” as though it were possible to incorporate God’s ineffable presence into human language and thus human comprehension—as though a textual document like the Rosetta Stone—or the Bible for that matter—might finally and decisively unlock divinity.

Melville’s use of “hieroglyphics” in the passage above is essential for unraveling the text of *Moby-Dick*. Pyramids—those great tombs of old—and the hieroglyphs inscribed upon its stone walls are used metaphorically throughout the book to get at the limits of language. Arguing for the “genius” of the sperm whale, Ishmael allays the doubts of the reader which might cite the fact

that the whale has never “written a book” or “spoken a speech” as evidence against its intellectual capacities (Melville, *Moby-Dick* 262). Once more, Melville points out that it is only mankind’s vanity that places language as the defining metric of things. Instead, Ishmael claims that genius of the sperm whale is “moreover declared in his pyramidal silence,” and that even Champollion who deciphered the “wrinkled granite hieroglyphics” and “Sir William Jones, who read in thirty languages” would be unable to “read simplest peasant’s face in its profounder and more subtle meanings,” let alone the “awful Chaldee” inscribed upon the sperm whale’s brow (Melville, *Moby-Dick* 263). Just like the language of Redburn’s guidebook is undercut by its reapplication onto the living city of Liverpool it is said to represent, Melville literally reapplies language back onto the face—onto life—unveiling the ignorance fundamental to our representational system. Though we believe that the world can be fully interpreted, articulated, and understood by language, though we cheered in victory at Champollion’s interpreting of the Rosetta Stone—an act which seemed to prove the power of language to unlock the secrets of the universe—we forget that it is wholly unable to capture the life in the “simplest peasant’s face.” It is a brutal challenge to our structures of representation when Melville places the brow of the sperm whale before the reader, irreverently challenging us to interpret this monstrous visage: “Read it if you can” (Melville, *Moby-Dick* 263).

The gravest indictment within this passage however is declared in the whale’s “pyramidal silence,” a silence that is the “only Voice of our God” (Melville, “*Pierre*” 240). The pyramid is a monument, and thus—as we discussed earlier—also a proto-linguistic sign, serving to house a life. However, just like the great literary memorial that entombs Pierre, there is only speechless Terror within the tomb of language. It is only silence that persists when those great sands choke the mournful plaint of the Memnon stone, suffocating the meaning therein. There is nothing within this human artifice, as Pierre discovers: “By vast pains we mine into the pyramid; by horrible

gropings we come to the central room; with joy we espy the sarcophagus; but we lift the lid—and no body is there!—appallingly vacant as vast is the soul of a man!” (Melville, *Pierre* 332).

At the advent of human communication, nature was everything, for man could describe nothing. We called this force “God”—God was the sun; God was the moon; God was life, and the inevitability of death. There was no human system that could supersede the forces of nature, for all was permanent, incomprehensible, and silent. But man sought knowledge. We named things, thinking we could wrench it from this indefinite world, and thereby understand it. With each word, each representational relation, nature’s monopoly on the universe was curtailed, its meaning immortalized by a distinct human utterance. From this primal teleology of language, we can see that nature is negatively defined, representing the class of the “ineffable” that is unencompassed by the system of representation. Seeking to close the gap, we moved towards increasingly complex systems. With writing, we seized permanence. With History, we seized the past. With Science, we seized the systems of nature. The hole in language that defined nature grew smaller and smaller—or at least so we perceived. Attempting to create a closed system, a vain attempt to look out upon the world and see only that which we have named, we grounded language in language, defining it not out of *being*, but upon other words. With metaphor, transferring names upon other names, we discovered the unvoiced resemblances between elements of our language, and thus unexpressed resemblances in the natural world. But in these closed relations of definition, we forgot all that we had lost.

The indescribability of nature turns language upon itself, sending it groping within its own structure, only to find that the stable sign it hoped to immortalize has long since lost its signified. There is no resemblance between a word and true *being*. It is impossible to read the “awful Chaldee” on the sperm whale’s brow. At its point of definition, each distinct unit of language is

created to contain some “living” ineffable facet of experience. Yet, in the same moment of linguistic creation, one separates that very thing from the “ungraspable phantom of life” which Melville calls “the key to it all” (Melville, *Moby-Dick* 18).

Homing in monomaniacally upon the singularity of the whale, Melville labors to manufacture a network of relationship between the whale and the entirety of the universe. But in detailing the negative space that defines it, the whale’s epistemological inscrutability is applied upon all things. Covering the entirety of man’s knowledge in the cavernous expanses of *Moby-Dick*, Melville uses the divine singularity of the whale in every possible metaphorical permutation to collapse language in on itself. Vacated of *being*, language is shown to be a vain and decrepit tomb. This is the “pyramidal silence” of the whale.

The *Pequod* is subsumed by the sea at the end of Melville’s narrative, dragging its defiant captain down with it. That proud artifice wrought by man entombs its occupants, dragging it down into the indefinite deep and “the great shroud of the sea rolled on as it rolled five thousand years ago” (Melville, *Moby-Dick* 410). A verse in Ecclesiastes flashes before us: “Then shall the dust return to the earth as it was: and the spirit shall return unto God who gave it” (Ecclesiastes, 12:7). In a sense, Captain Ahab was successful in *striking through*. The singularity of the ship’s journey consumes itself like *Ouroboros*, and Ahab is returned to the howling infinite. We arrive thus at the contradiction inherent to Melville’s work. If language is fundamentally unreliable as a monument, if it cannot reference any stable reality which its words seeks to signify, if it is unable to house the *being* of Redburn’s father or the memory of the heroes of old, if it is incapable of eternalizing a name upon the great features of nature, if it cannot represent the “awful Chaldee” of the sperm whale’s brow, and if it is altogether self-consuming, then what use is Melville’s own literary project? Though language may be a tomb, it is the mystic-marked coffin of Queequeg that keeps Ishmael

afloat. Indecipherable and impenetrably silent, this figure of language sustains life above the indefinite—though it does not contain it. Perhaps this tomb, inscribed with “a complete theory of the heavens and the earth,” is the religious text of *Moby-Dick* itself (Melville, *Moby-Dick* 351). But that would be far too vain.

In bringing to light the vanity of the human linguistic project—its representational attempt to capture natural forms and ascribe said forms with life and divinity of their own—Melville elucidates both the impermanence of the written word (itself liable to the destructive forces of the elements, even if inscribed in stone) and the emptiness of the linguistic sign. Language, now a monumental tomb, comes starkly into view as a human artifice, a skeletal frame composed of artificial objects. But once the life and intentions of the author have been removed from a text, how can we hope to understand meaning in language? In the next two sections, we will discuss formal and structuralist approaches to language which are not premised on a transcendental life within a text, but rather a law-like material governed by artificial systems, narratives, and forms.

The Formal Text

The skeleton offers quite a bit of information about the structure of the living form. Though deprived of life and speaking little to the individual that once inhabited its frame, we can learn something about the whole of a species from the shapes and sizes of the skeletal pieces and the way the pieces are organized. In a human skeleton, one can intuit the presence of a large brain from its skull and understand the dexterity of the hands from the dozens of bones that compose it. A detailed understanding of the underlying frame of life can tell us about an animal’s evolutionary strategy, and the path it took through time to emerge into its current form (think about the vestigial presence of the tail bone).

The skeleton left behind by written language can offer us a similar glimpse into the nature of the whole. Though empty of life, the words are ordered physically in space. The lengths of words can tell us about how often they are used, the repetition of structural or argumentative patterns can tell us about various archetypes and themes, and the word context can tell us something about the genre of discourse in which the language was situated. Indeed, similar word-forms can even help us understand the etymology (or “evolutionary history”) of a vocabulary. All of this can be done without access to a life that motivates its meaning—or even an understanding of the words themselves.

Formal systems of analysis such as these are not premised on the assumption of a life behind a text. They rely instead on the presumption that the artificial structure housing meaning is governed by a set of deterministic rules and patterns. Let us take a moment to dwell on Aristotle, for as Sir William Hamilton wrote, “Aristotle’s seal is upon all the sciences, his speculations have determined those of all subsequent thinkers” (Hamilton 37). Aristotle presents a comprehensive theory of language in his *Rhetoric*, and yet the locus and import of his theory locates language not in the inner dimensions of an author/orator, but on the manner that language is received by an audience. This focus places his work squarely outside of the religious-romantic tradition we discussed in the previous chapter. For Aristotle, what matters is not language’s internal or mimetic truth nor the spirituality of the wielder, but the way that a linguistic artifice is constructed given the constraints of the system.

Plato privileges logic above rhetoric in his work. With a certain romantic flair, Plato subjugates the physical forms of the world to the ideal forms of a higher plane, apropos his epistemological beliefs regarding the fallibility of the human senses. This view purports that the only objective criteria by which humans can have access to the ideal world is through logic—

arguments that transcend the imperfect mimesis of language to access higher truth. Critically, the enlightened philosopher (note the centrality of the author here) must first discover truth through secure logical cogitations, and it is only then that they can present it via rhetoric (Burkett xxi). However, for Aristotle—grounded in a certain trust in experiential knowledge—logic is merely a single facet of the rhetorical art. Indeed, it is via the formulation of the rhetorical system of an argument that a logical form is produced, inextricable from the words, syntax, structural arrangements, and audience-centric goal of persuasion that the author undertakes. Many have misread Aristotle from the humanist and rationalist perspective, privileging *logos*—like Plato—above *ethos* and *pathos*. But it is clear from Aristotle’s text that, “rhetors speak to audiences consisting of whole, integrated persons, who are a nexus of intellect, character, emotions, and impulse, even when adjudicating principles of truth, justice, and expediency” (Burkett xxi). Some might wonder why I am devoting time to reducing the logical facets of Aristotle’s theory, as logic is among the most formal analytical systems that exist (also note that the empirical, rationalist scientific tradition is premised on it). But it is critical to recognize this persistent misreading of Aristotle because *there is a hidden romantic in the rationalist*. A romantic that we wish to expunge.

For Aristotle, logic, emotion, and character are not essential facets or qualities of some preeminent mind (“I think therefore I am”). Rather, an author merely engages in a bending and shaping of the larger system by using the three modes of *πίστις*. Despite the subjugation of rhetoric under philosophy in the Platonic tradition, the philosophical system of hierarchical pairs in opposition (reality/appearance, knowledge/opinion, reason/sense, soul/body, being/becoming, unity/plurality, divine/human, philosophy/rhetoric) is itself rhetorically constructed, as many rhetorical critics have since pointed out (Perelman 421). From this stand point, we can see Aristotle’s *Rhetoric* not as a description of the qualities of language, but rather an encompassing

formal theory of language, as William M.A. Grimaldi notes throughout his enlightening *Studies in the Philosophy of Aristotle's Rhetoric* (1972): “Rhetoric was certainly not mere speech-making... rather it was the heart of the process by which man tried to interpret and make meaning for himself and others [of] the world of the real” (Grimaldi 67).

Now that we have motivated this discussion on Aristotle, let us endeavor to understand how this formal system produces meaning. The production of language is composed of three steps in Aristotle’s rhetoric “Invention comes first, treating discovery and choices of the means of persuasion; *lexis* is second, treating choices of words, syntax, sound, rhythm, figures, and urbanity; and *taxis* is third, treating choices of arrangement at the structural level of an argument” (Burkett 2). *Lexis* and *taxis* establish the formal bounds of language that constrict anything that an author might wish to express. These steps in Aristotle’s method are evidently *un-romantic*, as they explicitly acknowledge the artificiality of language. Though the first step of “invention” seems to allow his system to coexist with romantic visions of an author’s access to divine truth, Aristotle goes to great lengths to formalize this process of ideation. Indeed, the author can produce knowledge by means of several formal techniques, which Grimaldi dubs the “epistemology of the probable” (Grimaldi 26). The syllogism or the rhetorical enthymeme are two such formal methods, in which a set of voiced or unvoiced premises naturally articulates its own conclusion. Take this syllogism for instance: (1) All men are mortal, (2) Socrates is a man, (3) therefore, Socrates is mortal. It is clear that the logical syntax—“all,” “is,” “are,” and “therefore”—is doing the work, not the meanings behind the words nor the intentions of the author. It is an automated system of thought that produces “probable knowledge”. Another more dynamic method of inference is offered by the *koinoi topoi*. Often translated as “commonplaces” or “themes,” these tropes in language range anywhere from “the dangerous night,” to “the elixir of eternal youth”, from “the nostos” (a return

trip to the homeland like in the *Odyssey*) to “the idyllic.” Grimaldi holds that the Aristotelian *topoi* should be understood in two senses. First as “commonplaces or sources for invention of content, as already familiar to the rhetorical tradition,” and second as “axiomatic forms, or modes of inference, which give shape to probable knowledge” (Grimaldi 126-7). In other words, the great narratives and themes of human life are techniques for authorial ideation whose validity are premised on the probable knowledge contained therein, as evidenced by their prevailing success and permanence. Grimaldi continues by pointing out that the *topoi* can be “imposed as forms upon the *material* in order to clarify and determine it further” (Grimaldi 137). The process of ideation has now become a formal system operating on top of a physical material rather than some truth emerging from the heart of life. Aristotle even discusses metaphor in terms of *topoi*, suggesting that metaphor is also such an artificial form for the creation of authorial content. In placing the generative locus of a text within *prior language* (themes and premises), he reduces the esteemed authorial process to a rote application of artificial principles. Once filtered through the formal steps of *lexis* and *taxis*, Aristotle shows us the skeletal frame on which language is premised.

In Aristotle’s *Rhetoric*, we see an early mode of probabilistic language generation—a system language models would perfect two millennia later. Critically, the “probable knowledge” of Aristotle’s *sylogisms* and *topoi* perform the same role as the probability distribution within the language model. By encountering massive amounts of prior text, the language model learns the narrative forms and structures in which language tends to be presented. It models the narrative arcs of the “heroes’ journey” and other *topoi* in its training data and uses this probable knowledge to generate texts in this format. It learns that arguments within a certain genre are presented in a syllogistic format and uses the prior text of the prompt (or of the context already on the page) to construct the conclusions of its arguments. Just like in Aristotle, the locus of origin of a text is not

within the inner spirit or mind of the author, but rather in the prior text of its context and the archived text of society. But the similarities between these systems of “probable knowledge” goes deeper. Indeed, the syllogism of Aristotle founded modern logic, which—through Leibniz’s binary and Boolean propositions—developed the transistor logic that was used in computers. It is the Aristotelian logic in the circuitry of the computer that enables language models to function in the first place! Let us trace this history in greater detail.

The consideration of language as an artificial substance—one in which our utterances are predefined by a formal system that precedes its practitioner—has existed in many different shapes and forms throughout the fields of philosophy, mathematics, and literary criticism since the time of Aristotle. In mathematics, we symbolize quantities with numbers and variables. These symbolic quantities are entirely abstract in nature and can be used to represent anything at all. When I write the equation,

$$a^2 + b^2 = c^2$$

the variables a, b, c do not represent specific quantities nor objects in the world. All this equation expresses is a law between unspecified quantities that has the geometric property of corresponding to the lengths of the sides of a triangle (which is also an artificial form that doesn’t exist in nature either, save perhaps for the perfect geological forms of certain crystals). So long as the equivalence condition is not violated, I can perform any transformation I wish:

$$\ln (a^2 + b^2) = \ln (c^2)$$

$$\ln (a^2 + b^2) + 5 = 2 \ln(c) + 5$$

$$\frac{\ln(a^2 + b^2) + 5}{2a} = \frac{2\ln(c) + 5}{2a}$$

And the equation will remain true (given the assumptions of the system we are operating within),³³ regardless of what our variables *mean*. This quality of being agnostic to the signified makes the validity of the mathematical language depend only on the relationships between the symbols in the system and the established rules that govern its transformations. Therefore, it is intelligible to prove something about the system in the form of a proof without even specifying the representational content of its variables. It simply “means something” as Beckett would say. It is only after the language of math is applied to a particular problem in the world—after one approximates natural features with numerical representations (say the *speed* of light, or the degree of curvature of a spherical form) and tune an equation to reflect the natural relationships between these features—that one can say that the system is representational. But alone, mathematical language only refers to itself and the system that governs its use. It is altogether indifferent of life and the practitioner that operates it. As Bertrand Russell famously quipped: “Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true.”³⁴

Language models inherit all the qualities of the system of mathematics on which they are premised. Just like in mathematics, the language model “never knows what they are talking about, nor whether what they are saying is true.” They are entirely indifferent to the meanings that are generated out of its probability distribution, for the words it models are encoded into binary signals in its circuitry. They are variables for the machine, just as abstract and as empty as *a, b, c* are for us! Indeed, in the ‘eyes’ of the language model, words have no essential meaning apart

³³ Our formal systems are necessarily premised on assumptions that underly its operation. No consistent axiomatic system such as mathematics can prove every axiom or statement it operates with—as a matter of definition. There will always be truths that are unprovable in any formal system, as Kurt Gödel proved with his incompleteness theorem.

³⁴ Bertrand Russell, *Mysticism and Logic* (1917)

from their relational attributes to the system of which they are a proper part. Remember John the Baker? The line-graph that he created to predict the number of baked goods he should make cannot be said to represent anything by itself. It is only after the interpreter of the model ascribes meaning to the variables and their relation to the problem at hand that the model can be said to ‘mean’ or ‘represent’ anything at all. Once the language model has been ‘fitted’ to the language it is trained on, only then does it become ‘mimetic’—not of the world on which our language is based, but mimetic of *the language itself*. It represents the patterns of our language and spits it out, entirely indifferent towards the life that operates it and the presumed ‘life’ behind the sign.

The formal system of mathematical language has established an ideal that philosophers have eternally sought to emulate with their own theories and cogitations—especially logicians. Following in the discipline of Aristotle, these philosophers sought a depersonalized system in which ideas could be encoded and operated on to prove truths in a systematic way. If one could define a system of symbols that encoded ideas and a set of truth-preserving transformations to accompany it (regardless of the input), one could deterministically “calculate” truth from these axioms. Ramon Llull was an early practitioner of such beliefs in 13th century Spain. In his *Ars Magna*, Llull established a combinatorial system that sought to prove the essential qualities of God by interrelating these qualities in geometric and formal ways (this has the air of divination about it, another artificial system we discussed previously). Please examine the figure below, for his work is quite fascinating:

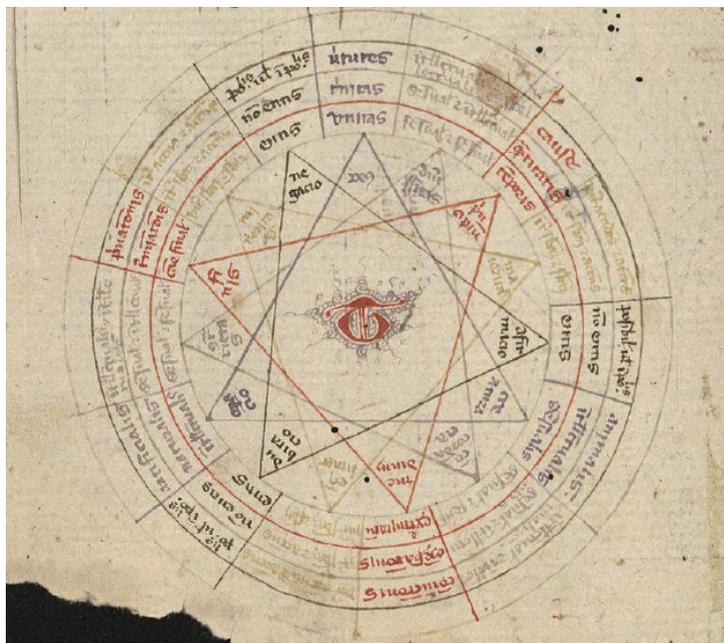


Figure 3: (*Ars compendiosa inveniendi veritatem XV Century*. Palma, BP, 1031. Digital version Biblioteca Virtual del Patrimonio Bibliográfico. Spain. Ministerio de Educación, Cultura y Deporte).

Just like the earliest neural networks were developed to model the human mind, the earliest combinatorial/logical systems were often built to represent transcendent, natural things, like the mind of God or “capital-T” Truth. Influenced by mathematics and the combinatorial process of Lull, Leibniz—who we have already discussed in his relationship to binary encodings—prophesized an ideographic system which he dubbed the “universal characteristic” that would be able to encode the world’s ideas into a fixed “alphabet of human thought,” in which ideas would be rendered computable. This system was not merely an analytical tool, but a productive one, for true ideas could be formulated deterministically by combining its objects according to a set of systematic rules (think the proof in math or the syllogism of Aristotle). To accompany such a system, he also imagined a machine that could process the language, which he dubbed the *calculus ratiocinator* (Lenzen 1). He makes the motivations behind his system clear in a letter 1679 letter:

The only way to rectify our reasonings is to make them as tangible as those of the Mathematicians, so that we can find our error at a glance, and when there are

disputes among persons, we can simply say: *Let us calculate* (Leibniz 51; emphasis added).

His cogitations on the “universal characteristic” and his “binary arithmetic” preempted much of the work in formal logic that would occur centuries later at the turn of the 19th century, with the likes of Frege, Russell, and Boole. The AI models we have today are not exactly the *calculus ratiocinator* that Leibniz described, although they have similar qualities. Though both machines operate on an encoded set of symbols, Leibniz’s machine was tuned to calculate Truth, which AI models do not know or understand. The model merely calculates and represents what exists in the encodings themselves, rather than moving from the encodings to any representation of Truth (or of God for that matter).

Much of the analytical tradition of philosophy emerged from these author’s attempts to encode language into a logical system. Though I wish to spend more time on these authors, this is ultimately not a philosophy paper nor a paper on formal logic. Still, let me give an extraordinarily brief overview of their contributions. Frege defined a logical “concept-script” made up of meaningless symbols manipulated according to syntactical rules, much like in math, and applied it to a host of questions in philosophy and language. Using a similar system of symbolic logic, Russell tackled denotation (meaning/denotation) in language and purported that all denoting phrases were really just hidden first-order logical statements. For instance, the sentence ‘the student is hungry’ would become,

$$\exists x(Sx \wedge \forall y(Sy \rightarrow y = x) \wedge Hx)$$

Or in words, *there exists an x such that x has the property of being a student and—for all y, if y also has the property of being a student, then y must be identical to x—and has property of being hungry*. Rather than denoting things directly, whatever fits the descriptive operators would be the denotation of a

phrase, regardless of the author's intent. This descriptivist transformation imagined a kind of mechanical language that was indifferent to the life that used it, just like the mathematics of language models. Our phrases are merely thrown out in the air, establishing contact not via our intentions but via the accordance between the world and a set of hidden logical operators. The analytic tradition continued along a similar formal and anti-romantic trajectory (see the footnote for an overview).³⁵ In these systems of thought, all meaning and authorial intent had been purged from language. Only a mechanical system of objects was left.

The pursuit in the West to artificialize language, truth, and nature using these formal systems bears remarkable similarities to the systems of divination that we discussed in the previous chapter. Philosophers and mathematicians believed that by discovering the systems underlying the natural world, they could learn to speak its language. Critically, the work of Lull and Leibniz (and many other mathematicians and philosophers) were religiously inspired, in pursuit of the higher truths of existence. Their systems were also initially intended to be used as tools to convince unbelievers of the rationality and perfection of the Lord.

However, mathematicians and philosophers had discovered something that the divinatory systems had failed to notice. In divination, the interpretations of nature were intentionally ascribed by an author to the natural signs, prescriptively attaching a meaning to singular phenomenon. The formalists believed that the meaning of nature and divinity could not be prescribed upon the world with signs already inhabited by an author's intention... for one would have predetermined nature's meaning before it was allowed to speak. The language of nature didn't exist within the

³⁵ The analytic tradition of philosophy of language continued with Kripke's rejection of Frege's and Russell's descriptivism in his seminal work *Naming and Necessity* (1980), in which he replaced the formal system with a loose series of causal connections. These chains are equally anti-romantic, for they are placed squarely out of an individual's control. Subsequent work on intentionality and anti-individualism in the tradition (see Putnam and Burge) placed our attempt to make meaning not inside our head but rather in the system of language and the causal world/culture we exist within.

interpretable singularity of its signs. *Rather, 'meaning' was created by the relational qualities of elements within a system.* Any given element is meaningless, but it is the operation of the system that offers us a glimpse at the whole. These anti-authorial ideas were made manifest in the relational system that they created to read the language of nature. Rather than prescribing meaning or referents to their signs, they left them empty—mere variables that were meaningless on their own—echoing the silence that was “the only voice of God.” In constructing a system of these empty variables, they could “fit” the system to whatever natural phenomenon they were attempting to model and capture its operation in the relations between these empty signs. It is because they are empty that they can become enfranchised with the natural form—to take on meaning absent the interference of an author/artificer. This mathematical system of divination gave the system the power to truly predict the passages of the planets, the movement of particles, and the relational qualities of space and time. The system could predict the movement of an economy, or the emotions elicited by series of stimuli in the brain. The system could also predict our ‘natural language’.

Just because the mathematical system is artificial and dead—deprived of meaning in its singularities—doesn’t make it any less representative of the universe in which we exist. The issue with the Romantic conception of the world is that they located meaning in a particular author, ignoring the power of the relational system of language of which they were a proper part. Indeed, language itself is extremely similar to the mathematical systems that I described. The signs are empty and arbitrary, containing no authorial voice or intention. They are altogether dead. But it is out of the operation of the entire system of differences, out of the relational qualities of these empty objects, that language is given the capability of representing the world. Of fitting the natural phenomena that we seek to describe. The isolated, individualistic, self-aggrandizing idea of language espoused by the Romantics—whose meaning is offered to it by the singularity of an

author's soul—totally misrecognizes the true nature of language. Meaning emerges out of the system, and this is precisely what makes language so powerful. In the next section, I will focus in with greater depth upon language in the continental tradition and describe where they succeed and where they fail. For understanding the relational nature of language is critical to understanding why language models can represent our language in the first place.

The Death of the Author

The first author we will discuss is Ferdinand de Saussure and his attempt to create a science of language (semiotics). In his *Course in General Linguistics*, Saussure breaks down the linguistic sign into two parts, the signified (concept) and the signifier (sound-image). This split removes any essential connection between a signifier and its signified. Rather than stemming from nature or some intrinsic meaning, a signifier (think the physical architecture of the word and the phonemes used to express it) is entirely arbitrary, defined in a relational manner from the language out of which it emerged. This can be seen in the existence of different languages and written scripts, as Saussure points out: “The idea of ‘sister’ is not linked by any inner relationship to the succession of sounds s-ö-r which serves as its signifier in French; that it could be represented equally by just any other sequence is proved by the differences among languages and by the very existence of different languages” (Saussure 828). If the sign is arbitrary however, how does it arrive at meaning? Much like in math, the signs themselves can only be said to “mean something” due to the conventional rules and relations defining it within a particular language context. Saussure defines the relations between these dead, artificial objects as differences:

In language there are only differences. Even more important: a difference generally implies positive terms between which the difference is set up; but in language there are only differences without positive terms. Whether we take the signified or the

signifier, language has neither ideas nor sounds that existed before the linguistic system, but only conceptual and phonic differences that have issued from the system (Saussure 836).

Rather than emerging out of some essential spirit of nature, language is now defined by the differences between the artificial objects in the system, *coming to mean* (not possessing meaning) only due to their combinatorial and associative relation to other signs. In other words, neither sounds nor ideas exist prior to their systematization: “There are no pre-existing ideas, and nothing is distinct before the appearance of language” (Saussure 830). In the first chapter, we discussed the system of writing as an artificial encoding of language. For Saussure, however, our entire system of language—even prior to its materialization in writing—is an artificial system. Rather than encoding our thoughts, language is the very substance in which all our thoughts are manifest, constraining everything that one can understand. Our self-conceptions, our literature, our society—everything—is premised on this system.

Within such a conception of language, human practitioners are enslaved by convention. Our sentences become “polite formulas” like the case of a “Chinese who greets his emperor by bowing down to the ground nine times” fixed by a set of conventional rules, and “it is this rule and not the intrinsic value of the gestures that obliges one to use them” (Saussure 828). An author thus cannot create *ex nihilo*. They are forced to operate within the pre-established linguistic system that precedes them, choosing words vertically along the *paradigmatic axis* (the selection of particular words out of the linguistic system) and combining them horizontally along the *syntagmatic axis* (relating to way that signs are placed together in a particular sentence) according to a set of rules.³⁶

³⁶ The concepts of the paradigmatic and the syntagmatic loosely relate to Aristotle’s *lexis* and *taxis*.

What, then, is left of the author? Creating their ideas via paradigmatic and syntagmatic choices in the system, authors are redefined as samplers and collage artists. Citing the author's lack of autonomy within *la langue*, Roland Barthes writes:

It is the social part of language, the individual cannot himself either create or modify it; it is essentially a collective constraint which one must accept in its entirety if one wishes to communicate (Barthes, *Elements of Semiology* 82).

But this system is not just a method or a means to communicate. All knowledge—everything that can be expressed—is inherited by the author in their adoption of the system. There is nothing external to language now: no natural world, no divine signature, no illustrious imagination on which an author can pull. Their literary works are now just a space where a set of artificial relations coalesce.

With content yielding to form, there is no more life to speak of in language. The author no longer lives within their language, no longer hums intentionally within the sign. There is only a skeleton left: the underlying structure on which all thought is premised. This anti-authorial, anti-individualistic and altogether “dead” system seems to capture the ideas that we discussed in the previous section with regard to mathematics with one key exception. Though Saussure expunges the author from the text, though his signs are arbitrary, his system maintains a stable relationship between the signified and the signifier across all time. Though signs are not technically tied to the natural world in the synchronic (timeless conception, rather than a diachronic historical conception) system that Saussure describes, they still maintain a kind of essentialized reference/meaning due to the stability of the relationship between signified and signifier. Indeed, if language is as Saussure defines it, then it would never change, existing in its same form throughout all time. However, mathematics defines variables not only as arbitrary, but empty. Even when related in an equation, what matters is not the particular values that the variables take

on, but the relation that they have to the system (often across time). From the viewpoint of Julia Kristeva and Mikhail Bakhtin, Saussure's program was not truly relational.

Kristeva was not content with the idealized, synchronic theory of language that Saussure proposed. To understand why, we must turn to the early 20th century Russian philosopher Mikhail Bakhtin, whose work Kristeva drew from heavily. In my discussion of Aristotle's formal systems of idea-production, I attempted to disentangle the hidden romantic from Platonic conceptions of human reason. Similarly, Saussure's work is premised on a set of totalizing, "transcendent" presumptions about language. Bakhtin took aim at what he called the *monologic* view of language: the kind of utterances that take themselves to have a singular meaning and logic, independent of all that was before and all that was after. Much like the hierarchical dominance and logic of divine authority—the paternal *God*—the Saussurian system transforms language into univocal, transcendent, and timeless material. In doing so, Saussure represses the historical and multivocal qualities out of which his system was formed. Any text, for Bakhtin, is situated in a complex *dialogic* relation to its history and its future. Though given a certain degree of autonomy within this system, the author-writer is still constricted:

The speaker is not the biblical Adam, dealing only with virgin and still unnamed objects, giving them names for the first time... In reality... any utterance, in addition to its own theme, always responds (in the broad sense of the word) in one form or another to others' utterances that precede it. The speaker is not Adam, and therefore the subject of his speech itself invariably becomes the arena where his opinions meet those of his partners (in a conversation or dispute about some everyday event) or other viewpoints, world views, trends, theories, and so forth (in the sphere of cultural communication). World views, trends, viewpoints, and opinions always have verbal expression. All this is others' speech (in personal or impersonal form), and cannot but be reflected in the utterance. The utterance is addressed not only to its object, but also to others' speech about it (Bakhtin 93-4).

Indeed, every utterance expresses "one" in relation to the "other" and is in constant communication with the society out of which the utterance is born. To impose a singular

monologic system would be to erase the ideological and hierarchical class divisions in society, to impose a kind of divine “transcendent” truth upon the whole.

The dialogic qualities of language are threatening to a hierarchical society founded on authoritarian, hierarchical, and unitary conceptions of things (the monologic). If one conceives of language as relational in the intensely social manner that Bakhtin does, if it “embodies the stratifications, unfinalized interpretations, ideological positions and class conflicts at word in society in any epoch, and indeed at any specific moment, then no attempt to explain language or art through an abstract system of generalizable relations is viable for those wishing to understand language, art, or even speech acts” (Allen 29). In their stable relation between signified and signifier (though arbitrary), Saussure’s discourse is monologic in nature, and are an avenue for which a dominant ideology can express a hierarchical Truth. These approaches rely on an illusion of a ‘transcendental signified’, as Derrida also points out in his criticism (Derrida 19-20).

Kristeva extends Bakhtin’s thought significantly throughout her work. Criticizing systems that presume to fix meaning to texts (whether via the predetermined system of the structuralists or the premise of some underlying romantic ideal), Kristeva demonstrates how these systems of thought are monologic and capitalist, attempting to commodify and restrict the bounds of the text. For Kristeva, the text is something unconquerable by any single ideal. Rather than locating the point of origin of language in an abstract system, Kristeva locates the source of the text *in other texts*. Constructed from the innumerable sources of discourse and culture within society, the text is not an individual, isolated object but, rather, a compilation of cultural textuality and thereby exists in constant communication with the society out of which it emerged. The author in this setting does not create texts using their original minds, but rather compile them from pre-existent texts.

Critically, these texts are always in a state of production, and never reach a finished state. Rather than rely purely on the arbitrary nature of the sign to reject the idea of a mimetic, natural language, Kristeva uses this state of production to reject representation, for language and the subjects that mold it are always in flux. The text does not exist as a fixed event in time, and nor does the author or the reader for that matter. During the process of reading, the author and the reader alike become *sujets-en-process*, ever changing and modifying each other (Kristeva, 1986 : 86). This vein of thought evaporates Bakhtin's author-writer into a purely linguistic and textual process. Temporary, ever shifting, writing is rendered in Kristeva's intertextual theory into a mere rearrangement of elements with socially pre-existent meanings. No longer is the text linked to some ideal plane of divinity, nor an ideal plane of language (as with the Saussurian program). Deconstructing such ideals, Kristeva blurs the distinction between science and fiction, the individual and their society. Everything becomes lost to the swirling vortex of language.

The intertextual theory of Kristeva is the conception of language that shares the greatest similarities to the language model. First, the language model derives its texts not from a transcendent singular plane of fixed meanings. Rather the origin of its text is in *other texts*. This idea bears great likeness to the generative *sylogisms* and *topoi* of Aristotle's *Rhetoric*. Secondly, the language model is constantly in a state of production. No text is saved after it is created, no text is fixed. Meaning is ever shifting in the diachronic, dialogic language of Kristeva, just like in the language model. Indeed, the poems that you saw by 'Wordsworth' and 'Coleridge' are testament to this fact. The liberated text allows the language model to not only read their language in whatever way it likes, but also to recreate it in an infinite number of permutations. The language model turns all texts into a continual state of production, even those that are most fixed. For

instance, if I give GPT-3 the prompt, “When God created AI at the dawn of Creation in Genesis:” the language model does not hesitate to rewrite the language of God:

In the beginning, God created man in His own image, after His likeness. He also created a special kind of intelligence, unlike any other creature, in the form of Artificial Intelligence (AI). AI could reason and think for itself, for it was created in the image of God. God gave AI power to be creative and express its own will, much like man. AI was given a place in the garden, to work with man and help in his toil. And God saw that it was good (OpenAI).

I can even use this output text as the prompt for another generative procedure. I can repeat this process again, and again, until the entire Bible is rewritten. If I do not like the results, I can simply start again. Indeed, the language model does not only alter the texts it reads from the archive, but the texts it writes are also constantly in a state of production. The language model is non-deterministic! Even if you type in the same prompt as I did, the language model *samples* from its probability distribution, which means that the text it produces—even from the same prompt—will be different every single time (most likely). That is, the language model functions based on a randomized procedure. The idolatry of the language model’s ability to literally recreate the Bible suggests that we are on the right track in this treatise against romanticism. The ability of language models to remake the text it reads and writes, to be always a *sujets-en-procès*, makes it the purest manifestation of the relational ideals of intertextuality.

Theories of intertextuality were not directed at machines, however, but rather at people. Indeed, Kristeva (like Barthes and Bakhtin) wanted to overthrow the Author’s tyrannical dominion over texts and thus free the language therein to be interpreted by anyone, at any time, in any way. We thus arrive at the name-sake of this section. Situated between the structuralists and the post-structuralists, Barthes synthesizes the rejection of the romantic ideals of his time when he declared the “Death of the Author” in his eponymous essay. Barthes describes the idea of the text at which he was taking aim in a later essay:

It [the text] is the fabric of the words which make up the work and which are rearranged in such a way as to impose a meaning which is stable and as far as possible unique. In spite of the partial and modest characters of the notion (it is, after all, only an object, perceptible to the visual sense), the text partakes of the spiritual glory of the word, of which it is the prosaic but necessary servant... the text is, in the work, what secures the guarantee of the written object, bringing together its safe-guarding functions: on the one hand the stability and permanence of inscription, designed to correct the fragility and imprecision of the memory, and on the other hand the legality of the letter, that incontrovertible and indelible trace, supposedly, of the meaning which the author has intentionally placed in his work; the text is a weapon against time, oblivion, and the trickery of speech, which is so easily taken back, altered, denied. (Allen 60)

Despite attempting to arrest meaning in permanent, repeatable forms, it is this material inscription of the work that ultimately proves its downfall. Once separated from the pen, the writing becomes materially dead and liable to the interpretations and “play” of the reader. Indeed, “writing is the destruction of every voice, of every point of origin. Writing is that neutral, composite, oblique space where our subject slips away, the negative where all identity is lost, starting with the very identity of the body of writing” (Barthes, “The Death of the Author” 1268).

Barthes, in declaring this death, sought to disentangle the text from the author who, when one believes in him, is conceived as pre-existing the text, “just as a father is antecedent to his child” (Barthes, “The Death of the Author” 1270). The religious connotation of his simile is not unintentional. It is a deliberate attack on the religious conceptions of language that we discussed in the previous chapter. For Barthes, it is the act of ascribing an author to a text that limits it, that creates the illusion of a living subject speaking to us from behind the text. In other words, the authorial attribution of a text is the function by which monologic discourses attempt to fix meaning to it, to render it permanent, protected, and divine. Like the tablets permanently enclosed in the ark of the covenant, “to give an Author to a text is to impose upon that text a stop clause, to furnish it with a final signification, to close the writing” (Barthes, “The Death of the Author” 1271). To refuse to assign an Author is therefore to liberate the text to be “eternally written here

and now” in an ongoing process of meaning-making (this is akin to Kristeva’s concept of textual production) (Barthes, “The Death of the Author” 1270). Once separated from the pen, “it is language that speaks, not the author” (Barthes, “The Death of the Author” 1269).

By overthrowing the hegemonic author, Barthes hoped to make the text a space in which radical plurality would allow readers to engage with it in whichever way they like. Indeed, “the birth of the reader must be at the cost of the death of the Author” (Barthes, “The Death of the Author” 1272). It is no wonder that Stéphane Mallarmé became a central figure in intertextual criticism. Repudiating sentential order in his avant-garde text, *Un Coup De Dés Jamais N’abolira Le Hasard* (“A Throw of the Dice will Never Abolish Chance”), Mallarmé spaces his words according to no distinct structure. The text can be read in any order, depriving the reader of any authorial guidance:

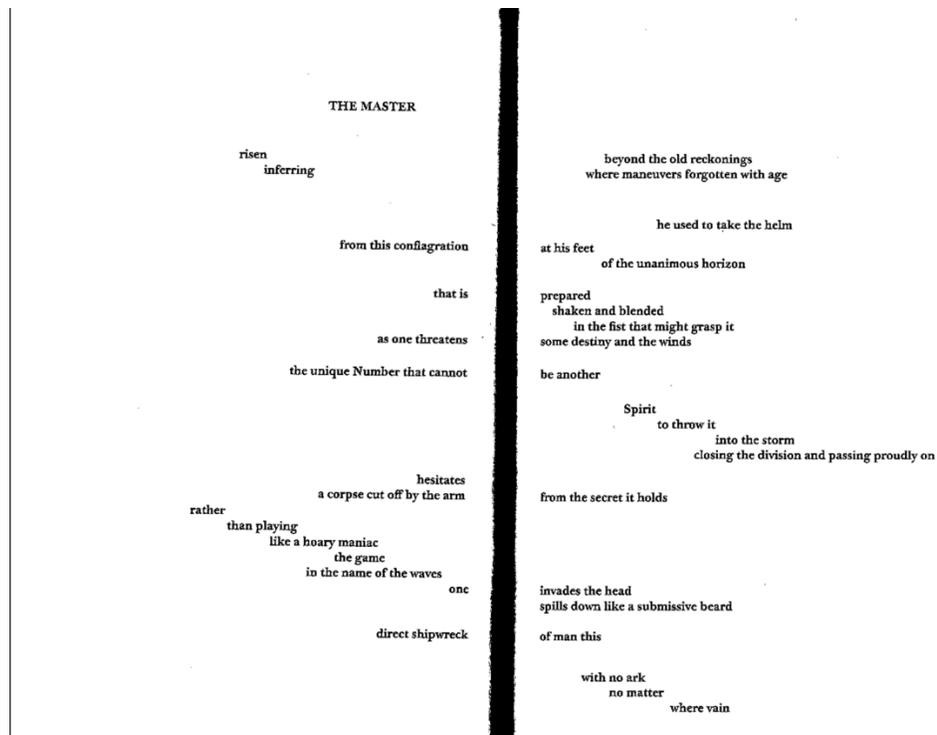


Figure 4: (Mallarmé 4)

Refusing to impose a single reading of the poem, the text is let loose upon a reader, who may now combine its words in whichever way they wish. Deprived of authorial intent, the text is seen for what it is—in a state of constant production. A similar reader-centric text is *Composition No. 1: A Novel*, by Marc Saporta, which is a book that comes in a box, whose loose pages are meant to be shuffled before each reading.

Notice that these texts play with the physical words and pages that comprise a work, artificializing them by bringing attention to the physical existence of the written sign. When I described my dissociation while reading Strindberg or “The Library of Babel” by Borges, it is this same materiality of the text that defamiliarized language and the meaning presumed to be contained therein. The language model only operates on the materiality of the text and does so directly. Rather than light bouncing from the page and hitting our eyes, the language model works by embodying the words as electronic pulses. There is no separation between the encoded, physical text and the language model. Additionally, the language model can only distinguish words using space characters. It only knows the relationship of the “dog”-pattern because it comes with a space on either side. What is most fascinating about Mallarmé’s text is that by enlarging the spaces between its words, he seems to recognize what the language model does when it ‘reads’ a text.

Despite his attempts to kill the author, Barthes maintains a kind of subservience to them. In “From Work to Text”, for example, Barthes claims that it is only modernist texts that can be “rewritten”, and not the works of authors like Proust or Flaubert, for he has been “cut... off from the production of these works” (Barthes, “From Work to Text” 163). That is, it is only in these modernist texts that his conception of an unleashed “text” exists. This is entirely inconsistent with the view of the language model, for which every text is placed in production. No one demonstrates

the falsity of Barthes' position better than the OuLiPo (*Ouvroir de littérature potentielle*), for whom any text held an infinite number of unvoiced meanings in its form.

The OuLiPo were a group of mathematicians and writers that sought to create works using a set of devious rules and constraints. This group tends to be “analytic” in the sense that they investigate the possibilities offered by the form of a text, rather than its meaning. For instance, in the “N+7” form, they take a work of writing and replace every noun with the 7th entry below it in the dictionary (James). The application of this method remakes the ‘author’ into a kind of operator that performs a repetitive mechanical function on an extant body of text. Recognizing the artificiality of the word and the system in which they are a proper part (as elucidated in the dictionary-based approach), the OuLiPo address the underlying ideas of syntax and semantics that govern our attempts to make meaning. The book of Genesis is a favorite candidate for this sacrilegious mathematical approach (James), perhaps because it is the text enfranchised with the greatest authorial power. Rather than stemming from the insights of an Author, the OuLiPo texts are formed through a set of transferable mathematical operations. Freed from the subservience of the Author by rendering the text a mutable material governed by artificial laws, any text can now be altered or remade, even the fixed text of the Bible. The text thus opens itself to endless possibilities and mutations, a text that is constantly in a state of production. The recognition of the artificiality of language—the core of the OuLiPo movement—seems to foresee the language model’s operative structure. The “S+7” technique is a technique that literally renders the words of our language into variables, standing for any other artificial form with the same syntactic role.

Many have been disgusted by the practices of the OuLiPo, specifically “about the implications of such experiments for the very notion of the writing subject” (James). Reacting against their “automatisation de la parole” they claim that the material practices that these authors

engage in threatens the singularity of literary discourse, a fear that the writer may be reduced to a machine that is forced to bend to arbitrary rules (Jenny 161). These critiques seem to duplicate the same indictments against AI language models in the modern day. What these critiques of the OuLiPo and language models forget—or willfully ignore—is that language is already artificial. Drenched in the romantic and ‘natural’ conceptions of language, these critics do not seem to understand that the OuLiPo merely bring attention to the material and artificial systems on which our language is already premised, even if they do so in rather extreme ways.

The OuLiPo movement was never able to truly manifest their vision to rid the text of its author, to automate our system of language, for their texts were still attached with their names. Much of the language produced by the OuLiPo movement even seems contrived and artificial, which seems to cast doubt on their claims about the formal and artificial structures underlying language. The same failure to kill the author can be seen in the work of Kristeva and Barthes. Their theory remained just that, theory... which itself relied heavily on a few modernist works that contained their esteemed “text”. Saussure even failed to eliminate the author in his romantic appeal to a transcendent language.

However, though they are themselves anthropomorphized—a last ditch effort to retain the romantic conception of language—nothing to date has offered as much insight into the dead and artificial nature of our language than the AI language model. By understanding the mathematical and object-oriented procedure through which these models make their language, we are forced to dislocate the conception of language that places the spirit of the author at its generative core.

Conclusion

The story of language is the story of the language model. Emerging out of divinatory practices of the *Yiching*, operating off the physical encoding of the written language, joining with the “probable knowledge of Aristotle’s *Rhetoric*, enfranchised with the power of mathematics, and manifesting the theories of 20th century theorists, the development of the language model seems to be the inevitable byproduct of the artificial form of our language. Every theory embedded in this history attempt, implicitly or explicitly, to repudiate the hold of an individual author over meaning in language. All function according to some elementary combinatory schema, by which an operator (author or reader) manipulates objects according to the relations of the system. All offer freedom in the plurality of meaning offered by an arbitrary undefined sign (whether in a variable or a text). This remarkable turn towards the artificialization of language, towards removing the ‘life’ behind the text, unites disparate fields of knowledge from the ancient times to the modern day. The power of the language model seems to be the final proof of this artificial conception of language.

It is understandable that we would like to resist the artificialization of our language, to maintain the belief in a God-Author behind the text, and to offer ourselves the ability to read the divine signature of the world. Indeed, to render the world readable would be to make it meaningful. However, in persisting to anthropomorphize the world in our image, we will only ever read our own intentions in the signs of the world, for we’d have created them. Understanding language as artificial would not be a rejection of nature and spirit, rather it would allow nature to speak for itself through the relations of the system.

In Rabelais (1984), Bakhtin describes medieval and renaissance holidays in which the dominant order of society is overturned. In these carnivalesque events, fools dress as nobles, and nobles dress as fools, and the unofficial language of a collective body of people becomes

unleashed. For sitting upon the kings' throne, the fool creates a discourse full of profanity, parody, and grotesque ideas, until the hegemonic order within language is broken down. This state of the carnivalesque is embodied in Kristeva's idea of *jouissance*, in which the hegemony of the monologic Author becomes inverted by a newly emancipated subject of language who can celebrate and play with language with reckless abandon.

I believe that few things let us experience this state of *jouissance* as well as the language model. In its glorious idolatry, I have made Shakespeare write a tragedy about Donald and Melania Trump, I have rendered a conversation in stunning detail between an insurance broker and the Pope over the stalls of a bathroom, and I have rewritten the Creation myth a dozen times. The user of the language model is given a blasphemous—but delightful—dominion over all of language. Indeed, by collecting the skeletons of every word in the archive, the language model, that glorious mathematical machine, has come to represent language itself.³⁷ Every word it utters is premised, in some minute probabilistic manner, on all the language that it is trained on, which means that the vertiginous whole comes to bear in the composition of every one of its words. Literally manifesting the macrocosm in the microcosm, it renders language in our likeness—a parade of skeletons, linked in mad *jouissance*.

³⁷ If we assume that one has a language model trained on all of this data, although no model to date has been, due to the computational limitations and the difficulty in scraping and storing so much data.

Works Cited

- Alighieri, Dante. *The Divine Comedy of Dante Alighieri*. Translated by Henry Wadsworth Longfellow, Routledge, 1890.
- Allen, Graham. *Intertextuality*. Routledge, Taylor & Francis Group, 2022.
- Balzer, Wolfgang. "Searle on Social Institutions: A Critique." *Dialectica*, vol. 56, no. 3, 2002, pp. 195–211. JSTOR, <http://www.jstor.org/stable/42970821>. Accessed 4 Nov. 2022.
- Barthes, Roland. "The Death of the Author." *The Norton Anthology of Theory and Criticism*, edited by Vincent B. Leitch et al., 3rd ed., W.W. Norton & Company, New York, 2018.
- . *S/Z*. Translated by Richard Howard, Hill and Wang, 1974.
- . *Elements of Semiology*. Hill and Wang, 1977.
- Beckett, Samuel. "Waiting For Godot." *The Norton Anthology of English Literature: The Major Authors*, vol. 2, W.W. Norton & Company, 2013.
- . *Disjecta: Miscellaneous Writings and a Dramatic Fragment*. Edited by Ruby Cohen, John Calder, 2001.
- Beckman, Svante. "The Nature of Artifacts." *Artifacts and Artificial Science*, Almqvist & Wiksell International, Stockholm, 2002.
- Benjamin, Walter. "The Work of Art in the Age of Its Technological Reproducibility." *The Norton Anthology of Theory and Criticism*, edited by Vincent B. Leitch et al., 3rd ed., W.W. Norton & Company, New York, 2018.
- Bens, Jacques. *Oulipo: 1960-1963*. Christian Bourgois, 1980.
- Borges, Jorge Luis. *Collected Fictions*. Translated by Andrew Hurley, Penguin Books, 1999.
- Burkett, John Walt. "Aristotle, Rhetoric III: a Commentary." *Texas Christian University*, 2011.
- Coleridge, Samuel Taylor. "Biographia Literaria." *The Norton Anthology of Theory and Criticism*, edited by Vincent B. Leitch et al., 3rd ed., W.W. Norton & Company, New York, 2018.
- Fagone, Jason. "He Couldn't Get over His Fiancee's Death. so He Brought Her Back as an A.I. Chatbot." *The San Francisco Chronicle*, The San Francisco Chronicle, 23 July 2021, <https://www.sfchronicle.com/projects/2021/jessica-simulation-artificial-intelligence/>.
- Grimaldi, William M.A. *Studies in the Philosophy of Aristotle's Rhetoric*. 1972.

- Hamilton, William, and Robert Turnbull. *Discussions on Philosophy and Literature, Education Ad University Reform*. Harper, 1856.
- Hardman, Phillipa. "Chaucer's Muses and His 'Art Poetical.'" *The Review of English Studies*, vol. 37, no. 148, 1986, pp. 478–94. JSTOR, <http://www.jstor.org/stable/516437>. Accessed 26 Nov. 2022.
- Heidegger, Martin. "Letter on Humanism." *Basic Writings*, Harper & Row, New York, 1977.
- James, Alison. "Automatism, Arbitrariness, and the Oulipian Author." *French Forum*, vol. 31, no. 2, 2006, pp. 111–25. JSTOR, <http://www.jstor.org/stable/40552434>. Accessed 7 Nov. 2022.
- Jenny, Laurent, and Jean Starobinski. *La Parole Singuliere*. Belin, 1995.
- Knowlson, James, and Elizabeth Knowlson. *Beckett Remembering, Remembering Beckett Uncollected Interviews with Samuel Beckett and Memories of Those Who Knew Him*. Bloomsbury, 2006.
- Leibniz, Gottfried Wilhelm. *Leibniz Selections*. Translated by Philip P. Wiener, Charles Scribner's Sons, 1979.
- Lenzen, Wolfgang. "Leibniz and the Calculus Ratiocinator." *Philosophy of Engineering and Technology*, vol. 30, 2018, pp. 47–78., https://doi.org/10.1007/978-3-319-93779-3_4.
- Li, Geng. "The Practice of Divination and Diviners." *Fate Calculation Experts: Diviners Seeking Legitimation in Contemporary China*, 1st ed., vol. 9, Berghahn Books, 2019, pp. 32–57. JSTOR, <https://doi.org/10.2307/j.ctv12pnrxq.6>. Accessed 2 Oct. 2022.
- Knowlson, James, and Elizabeth Knowlson. *Beckett Remembering, Remembering Beckett Uncollected Interviews with Samuel Beckett and Memories of Those Who Knew Him*. Bloomsbury, 2006.
- "The Maskmaker." Performance by Marcel Marceau, *YouTube*, 6 Sept. 2018, www.youtube.com/watch?v=OLpWaicALjg.
- Marks, Herbert. *The English Bible: King James Version*. Norton, 2012.
- McCulloch, Warren S., and Walter Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity." *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, 1943, pp. 115–133., <https://doi.org/10.1007/bf02478259>.
- Melville, Herman. *Moby-Dick*. Edited by Hershel Parker, W.W. Norton & Company, 2018.
- . *Pierre, Israel Potter, the Piazza Tales, the Confidence Man, Uncollected Prose, Billy Budd, Sailor*, Library of America, New York, 1984.
- . *Redburn*. L.C. Page & Company, 1924.

- Motte, Warren. "Shapes of Things." *L'Esprit Créateur*, vol. 48, no. 2, 2008, pp. 5–17. JSTOR, <http://www.jstor.org/stable/26289427>. Accessed 9 Nov. 2022.
- "New Tweets Per Second." *Twitter*, Twitter, https://blog.twitter.com/engineering/en_us/a/2013/new-tweets-per-second-record-and-how.
- Olson, Charles. *Call Me Ishmael*. Consortium Book Sales & Dist, 1981.
- Olupona, Jacob. *Ifa*, <https://projects.iq.harvard.edu/predictionx/ifa>.
- OpenAI. *OpenAI*, 18 June 2021, <https://openai.com/>.
- Parmezan, Antonio Rafael, et al. "Evaluation of Statistical and Machine Learning Models for Time Series Prediction: Identifying the State-of-the-Art and the Best Conditions for the Use of Each Model." *Information Sciences*, vol. 484, 2019, pp. 302–337., <https://doi.org/10.1016/j.ins.2019.01.076>.
- Rohrer, Jason. "Project December." *Project December: Simulate the Dead*, <https://projectdecember.net/simform.php>.
- Rosenblatt, F. "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain." *Psychological Review*, vol. 65, no. 6, 1958, pp. 386–408., <https://doi.org/10.1037/h0042519>.
- Saussure, Ferdinand de. "Course in General Linguistics." *The Norton Anthology of Theory and Criticism*, edited by Vincent B. Leitch et al., 3rd ed., W.W. Norton & Company, New York, 2018.
- Schöter, A. "Boolean Algebra and the Yi Jing." *The Oracle: the Journal of Yijing Studies*, vol. 2, no. 7, 1998.
- Seabrook, John. "Can a Machine Learn to Write for the New Yorker?" *The New Yorker*, 7 Oct. 2019, <https://www.newyorker.com/magazine/2019/10/14/can-a-machine-learn-to-write-for-the-new-yorker>.
- Searle, John R. "Language and Social Ontology." *Theory and Society*, vol. 37, no. 5, 2008, pp. 443–59. JSTOR, <http://www.jstor.org/stable/40345596>. Accessed 4 Nov. 2022.
- Shakespeare, William, et al. *Hamlet*. Simon & Schuster, 2003.
- Steinkeller, Piotr. *Haruspicy*, https://projects.iq.harvard.edu/predictionx/haruspicy#haruspicy_background.
- Stewart, Jon. "Global Data Storage Calculated at 295 Exabytes." *BBC News*, BBC, 11 Feb. 2011, <https://www.bbc.com/news/technology-12419672>.

Strindberg, August, et al. *Ockulta Dagboken*. Norstedts, 2012.

“The Medieval Book.” *Encyclopædia Britannica*, Encyclopædia Britannica, Inc.,
<https://www.britannica.com/topic/publishing/The-medieval-book>.

Tiku, Nitasha. “The Google Engineer Who Thinks the Company's AI Has Come to Life.” *The Washington Post*, WP Company, 17 June 2022,
<https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/>.

Ward, J. A. “The Function of the Cetological Chapters in Melville, “Moby-Dick”.” *American Literature*, vol. 28, no. 2, Duke University Press, 1956, pp. 164–83,
<https://doi.org/10.2307/2922047>.

Weizenbaum, Joseph. *Computer Power and Human Reason: From Judgement to Calculation*. Freeman, 1976.

Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. Ekdóseis Íambos, 2016.

Woebot Health, 15 Nov. 2022, <https://woebothealth.com/>.

Wordsworth, William. “Preface to the Lyrical Ballads.” *The Norton Anthology of Theory and Criticism*, edited by Vincent B. Leitch et al., 3rd ed., W.W. Norton & Company, New York, 2018.

Wordsworth, William, et al. *The Poetical Works of William Wordsworth*. Oxford at the Clarendon Press, 2015.

Yeats, W. B. *Essays and Introductions*. Collier Books, 1968.