

## Creative Minds Like Ours?

### Large Language Models and the “How” of Human Linguistic Creativity

By Vincent J. Carchidi

#### Abstract

The creative aspect of language (CALU) is the stimulus-free, unbounded, yet appropriate use of language by human beings. This use of language is entirely ordinary. While this concept—which amounts to a set of observations and descriptions about human linguistic behavior—is most frequently associated with the work of Noam Chomsky and other generative linguists, it is a concept whose acceptance is independent of commitment to any particular linguistic theory. This is but one reason why CALU’s implications for the study of both human and computational creativity have been unduly neglected. In contrast to this neglect, this paper highlights how this distinctive ability is responsible for the human coordination of language as a resource for creativity—it represents the *how* of human creativity, rather than *what* is produced by humans. Drawing from recent work done at the intersection of human and computational creativity using work from cognitive science and philosophy (Franceschelli and Musolesi, 2023), this paper seeks to remedy CALU’s absence. It does so by articulating and defending this concept and putting it forward as a baseline account of how human beings engage in creative behaviors. A test of whether state-of-the-art generative Large Language Models (LLMs) exhibit this ability, according to the three criteria of stimulus-freedom, unboundedness, and appropriateness to circumstance, is conducted. Remarks at the end are given on the relationship between generative linguistic theory and CALU and how this may shed some light on future developments in computational creativity.

**Keywords:** Artificial Intelligence; Large Language Models; Computational Creativity; Creative Aspect of Language Use; Generative Linguistics; Cognitive Science

## 1 Introduction

The creative aspect of language use, or CALU, is a neglected yet critically relevant feature of human creativity. Typically associated with generative linguistics, and more specifically with the work of Noam Chomsky, CALU describes the *stimulus-free, unbounded, yet appropriate* use of language by humans (McGilvray, 2001, 6-13; Asoulin, 2013, 228-232). While often confused with abstract knowledge of language, or linguistic competence—due, in no small part, to a lack of clarity in Chomsky’s early writings—this feature of linguistic performance, or behavior, is entirely ordinary, reflecting everyday uses of language. Such confusion has contributed to its absence in the rich study of human creativity aided by cognitive science and philosophy. As a result, CALU has similarly been absent in the study of *computational* creativity and what it would mean for artificial intelligence (AI) to exhibit human-like creativity.

This paper seeks to remedy this absence. In doing so, a defense of CALU is offered while simultaneously locating it within the relevant literature on human and computational creativity. It builds on recent work done at the intersection of creativity studies and large language models (LLMs) by Giorgio Franceschelli and Mirco Musolesi (2023). There, they write: “Creative-person qualities in Generative AI might eventually be the ultimate step in achieving human-like intelligence” (Franceschelli and Musolesi, 2023, 6). Disagreements about the role of generative AI *specifically* in achieving some higher form of AI aside, they are on to something about creativity’s role in this endeavor. They are right to recognize that the ability to exhibit human-like creativity is a matter of technical and practical importance for AI and that cognitive science and philosophy shed light on understanding this intersection of research.

While Franceschelli and Musolesi do not explore CALU, they do inaugurate a conversation on the creativity of LLMs and how this relates to the technical development and potential uses of these systems that is grounded in cognitive science and philosophy. They make prominent use of cognitive scientist Margaret Boden’s (2004) conception of creativity in their assessment of LLMs, while also expanding their scope to the full breadth of possible factors entering into creativity—not just *what* a creative product is, but *how* it is generated. We will see how CALU relates, specifically, to the *how* of the creative *process* and bring this into interaction with broader conceptions of creativity.

This paper attempts to explicate CALU in a largely theory-neutral manner. Attention is, of course, paid during CALU’s explication to the primary source material where generativists like Chomsky believe CALU was first articulated—in the work of René Descartes (1637), where he drew attention to the use of language as a test for the existence of minds like ours in contrast to automata and animals. There is no necessary commitment, however, to a specific linguistic theory, including theories within generative linguistics, in recognizing that CALU exists. This theory-neutral angle may sound odd given the concept’s association with Chomskyan linguistics, but this characterization of human linguistic performance is not dependent on any one particular linguistic theory and is thus widely applicable across conceptions of human and computational creativity.

The association of CALU with the study of human and computational creativity is not merely a convenient foil for further argument. More than this, echoing remarks above by Franceschelli and Musolesi, creativity in humans and machines reflects technical matters related to the natures and architectures of human minds and AI systems. This paper thus seeks to establish CALU as part of a baseline for creativity in human beings. Specifically, it seeks to show how CALU

enables the *how* of creativity—the ways in which a creative product is generated—while interacting fruitfully with timely matters of AI, namely LLMs.

Here is how this is done: a survey of the literature on human and computational creativity is provided, with particular emphasis on Boden’s conception of creativity and how Franceschelli and Musolesi use this to assess the creative potential of state-of-the-art LLMs. From this, we extract points of focus highlighted by these researchers relevant to CALU, thus proceeding to this concept’s extensive explication. This explication is complete with a defense of the concept as well as reasons for its absence in the relevant literature and objections to the concept, to which all are responded. We then dive into the computational implications of CALU, exploring how this ability manifests in humans, and conducting a straightforward test of whether LLMs exhibit this behavior according to its three criteria of stimulus-freedom, unboundedness, and appropriateness. Closing remarks are made in reference to linguistic theory—a topic largely avoided until the end—and how CALU relates to technical and practical matters in the future of human and computational creativity.

## 2 Literature Review

Creativity studies encompass both philosophy and science, and several distinctions and concepts are critical to locating specific frameworks within the field. Among these is the distinction between *human* creativity and *computational* creativity. Whereas the former refers to the study of the “characteristics and cognitive behavior of creative people and the environment or situations in which creativity is facilitated,” computational creativity “is often expressed in the formal language of search spaces and algorithms” while inspired by human creativity along the way (Maher, 2006, 22). It is the distinction between “understanding *human* creativity” and “trying to produce machine creativity...in which the computer at least *appears* to be creative to some degree” (Boden, 2004, 1).

Several criteria have been proposed in this vein as metrics to *evaluate* whether an idea or artifact (a product) is creative. The representative list of criteria suggested by scholars includes the nature of specific artifacts—say, a poem considered in complete isolation and on its characteristics alone—as well as the mindsets, intentions, and actions of those responsible for producing them and the social contexts in which they are produced. Franceschelli and Musolesi (2023, 6), drawing from Mel Rhodes (1961, 307-310), offer a useful overview of these criteria. They divide them into *product* and *process*—the creative artifact or outcome itself and the role of motivation, thinking, communicating, and so on in generating it, respectively—and *person* and *press*—the relationships between the potentially creative product with the individual that produced it as well as the social and historical context of its origin. This echoes Graeme Ritchie’s warning that “any formal definition of creativity must be based on its ordinary usage; that is, it must be *natural* and it must be *based on human behaviour*” (Ritchie, 2007, 69). Without using a conception of creativity derived from its ordinary human uses, the danger of circular argument looms (Ritchie, 2007, 69-70).

Consider, as a prominent example through which to explore these criteria, Boden’s definition of creativity from her major work, *The Creative Mind*: creativity is “the ability to come up with ideas or artefacts that are *new, surprising and valuable*” (Boden, 2004, 1). *Ideas* refer to concepts, poems, musical compositions, and the like, whereas *artefacts* refer to paintings,

sculptures, pottery, and so on (Boden, 2004, 1). Here we have the three criteria of *novelty*, *surprise*, and *value*.<sup>1</sup>

While Boden does, as we see below, develop a taxonomy of ways of being creative, her definition highlights that some scholars place emphasis not only on the creative idea or artifact and the impact it has on humans, but also on how it is generated and the relationship it bears to the creator—*product*, *process*, *person*, and *press*. Berys Gaut argues, subdividing this further, that “*how* the original and valuable product is made plays an essential role in determining whether the act of making it is creative...the making must involve *flair* by the maker to rule out” products generated “by chance or by mechanical procedure” (Gaut, 2003, 270). Gaut thus defines a process as creative “when it is the producing of something valuable and original by flair” and artifacts as creative “when they are original, valuable, and produced by flair. Originality, value and flair are the vital ingredients in creative making” (Gaut, 2003, 271). *Process*, *product*, *press*, and *person* all matter to Gaut. Luciano Floridi and Massimo Chiriatti echo their support of an expanded notion of creativity in their affirmation that “it is not *what* is achieved but *how* it is achieved that matters” (Floridi and Chiriatti, 2020, 687).

Finally, the distinction between what are called *P-creativity* and *H-creativity* is important to highlight, as they relate importantly to computational creativity. In frameworks like Boden’s, *novelty* can have two distinct meanings. One refers to the extent to which “the produced item [is] dissimilar to existing samples of its genre” (Ritchie, 2007, 72). The other refers to the “property of not being in existence before,” which itself can mean relative to the individual or relative to all of history (Francescella and Musolesi, 2023, 4). P-creativity and H-creativity capture these differences: P-creativity “involves coming up with a surprising, valuable idea that’s new to the person who comes up with it” and H-creativity “means that...no one else has had [the idea] before: it has arisen for the first time in human history” (Boden, 2004, 2). This distinction is often taken seriously in the study of computational creativity because it is P-creativity—the process of creativity relative to an agent’s “initial state of knowledge”—that is relevant to the “mechanisms of creation” and the goal of modeling it within computer programs (Ritchie, 2007, 72).

Where does CALU fit into all this? As we will see, the criteria necessary for the existence of CALU in human behavior are dependent primarily on observable features and descriptions of linguistic behavior that can be identified largely independently of any specific linguistic theory. This is consistent with Ritchie’s above emphasis on a definition of creativity (formal, in his case), being rooted in *ordinary* human behavior.<sup>2</sup> In addition, CALU falls under the *how* dimension of creativity and does not, directly at least, inform scholars of the characteristics of *what* comprises a creative artifact. Finally, while CALU may fall into either P-creativity or H-creativity (linguistic utterances may be not only novel to an individual but to all of human history), the relevance of this particular distinction is more associated with theories about the human language capacity specifically, rather than CALU itself, as we see in concluding remarks.

## 2.1 Boden’s Taxonomy of Creativity

Boden’s taxonomy of creativity—which includes *combinatorial* creativity, *exploratory* creativity, and *transformational* creativity—represents, for her, the three major ways in which creative

<sup>1</sup> Boden (2004, 2-3), to be sure, considers surprise to be a complement of novelty. More on this below.

<sup>2</sup> Whether these features are “empirically observable,” in a literal sense, is a matter we explore below.

ideas can be produced. Each of them possesses subtle relationships with her criteria of novelty, surprise, and value. The taxonomy is illustrated in Table 1 below:

<b>Forms of Creativity</b>	<b>Description</b>	<b>Examples</b>
<b>Transformational Creativity</b>	<p>Transformational creativity is the height of this ability, in which “the space or style itself is transformed by altering (or dropping) one or more of its defining dimensions. As a result, ideas can now be generated that simply <i>could not</i> have been generated before the change.” The “initial rules of the search space are altered” (Boden, 2009, 25).</p> <p>The resulting ideas alter the culturally accepted style of thinking, thereby making new ideas “more surprising—even shocking” (Boden, 2009, 25).</p>	<p>The rarest form of creativity, transformation occurs when ideas that are <i>fundamentally different</i> than past ideas are introduced, including in the artistic work of figures like Vincent van Gogh or selective achievements by scientists like Crick and Watson (Boden, 2009, 25).</p>
<b>Exploratory Creativity</b>	<p>Exploratory creativity involves moving through a culturally accepted conceptual space enabled and restricted by generative rules. It occurs when “the person moves through the space, exploring it to find out what’s there (including previously unvisited locations)—and, in the most interesting cases, to discover both the potential and limits of the space in question” (Boden, 2009, 25).</p>	<p>Examples of exploratory creativity include building scientific theories within an established conceptual space—that is, within accepted parameters that fit a particular “style of thinking” (Boden, 2009, 25).</p>
<b>Combinatorial Creativity</b>	<p>Combinatorial creativity “produces unfamiliar combinations of familiar ideas, and it works by making associations between ideas that were previously only indirectly linked” (Boden, 2009, 24).</p>	<p>Examples of combinatorial creativity include poetic imagery, artistic collages, and mimicry, among others.</p> <p><i>Analogy</i> falls under this form of creativity as it “exploits shared conceptual structure and is widely used</p>

		in science as well as art” (Boden, 2009, 24).
--	--	--

Table 1: Boden’s Taxonomy of Creative Ideas (see Boden, 2004, 3-6; Boden, 2009, 24-25 for concise overviews of Table 1’s information)

In brief: combinatorial creativity associates existing ideas with one another in unfamiliar ways; exploratory creativity builds new ideas within a conceptual space undergirded by a system of generative rules accepted by one’s society; and transformational creativity builds new ideas by radically altering one or more of the generative rules in the conceptual space accepted within one’s society.

Boden is careful to clarify that combinatorial creativity, while occupying the lowest rung, “is not to be sneezed at,” representing such “renowned” accomplishments as the works of Kurt Schwitters and Shakespeare “which depend not on stylistic transformations but on associative processes for their origination (and their interpretation, too)” (Boden, 2009, 25). Exploratory creativity, similarly, is what most artists and scientists—even those who very occasionally engage in transformative work—engage in throughout their careers (Boden, 2009, 25).

Boden’s taxonomy also makes significant use of the notion of *surprise*, though sometimes tucking surprise implicitly into the generation of novel ideas.<sup>3</sup> Francescelli and Musolesi observe how “[t]hese three different forms of creativity involve surprise at increasing levels of abstraction” (Francescelli and Musolesi, 2023, 4). The transformation of a conceptual space, the highest form of creativity, can lead to “shocking” new ideas (Boden, 2009, 25).

The emphasis on surprise in the study of machine intelligence traces back to the origin of modern AI. Summarizing a variant of mathematician Ada Lovelace’s argument that no ‘Analytical Engine’ (i.e., a digital computer) could possess powers of originality,<sup>4</sup> Alan Turing wrote: “a machine can never ‘take us by ‘surprise’” because they ““never do anything really new”” (Turing, 1950, 450).<sup>5</sup> The relation between novel ideas and surprise in Boden’s taxonomy, then, is one with historical roots.

## 2.2 LLMs and Combinatorial Creativity

In their application of Boden’s taxonomy to LLMs, Franceschelli and Musolesi (2023, 4-5) argue that autoregressive<sup>6</sup> LLMs fine-tuned via Reinforcement Learning with Human Feedback (RLHF) *minimize*, by design, the likelihood of surprise induced by their outputs. The result of such a training technique, they argue, is that “its strict alignment to very careful and pre-designed human responses leads to the generation of text that might be considered *banal*” (Franceschelli and Musolesi, 2023, 5). They thus argue that such systems may be capable of *combinatorial* creativity given the right prompting techniques and the sometimes-useful results this can yield,

<sup>3</sup> Boden (2009), for example, explicitly invokes surprise in her taxonomy, though she defines creativity as “the ability to generate novel, and valuable, ideas” (Boden, 2009, 24), lacking explicit emphasis on surprise as she does in *The Creative Mind* (Boden, 2004, 1).

<sup>4</sup> On this argument, see Boden (2004, 16-21).

<sup>5</sup> Turing, in contrast to Lovelace, reported being surprised by machines “with great frequency” (Turing, 1950, 450).

<sup>6</sup> “Autoregressive” simply refers to a general process in machine learning in which a system uses past values to predict future values. In LLM-powered chatbots, past tokens given as input, representing words or parts of words, are used to predict future tokens, which the end-user sees as words in natural language.

but transformational creativity is off-limits to LLMs that operate in a fundamentally probabilistic manner.

For Franceschelli and Musolesi (2023, 4-6), then, both *what* is achieved and *how* it is achieved matters, drawing attention to previously mentioned notions of individual flair and motivation in addition to self-evaluation. They pinpoint the reliance on end-users to *prompt* the LLM as a mechanism that reduces surprise and therefore novelty, even if sometimes proving useful (i.e., valuable), meaning the system is constrained to the lower levels of abstraction in Boden's taxonomy. They relate this to the importance of motivation in creativity, and "intrinsic interest" in a creative activity that leads humans to "produce different response possibilities and to internally test them in order to select the most appropriate." However, LLMs simply "do not contain such self-feedback...they stop at the first stage of creative learning, i.e., imitation, not implementing the remaining ones, i.e., exploration and intentional deviation from conventions" (Franceschelli and Musolesi, 2023, 5).

These limitations of LLMs are relevant, of course, to the technical future of such systems, particularly if researchers wish to construct systems capable of higher levels of creativity (according to Boden's taxonomy, anyway). Notice, however, that they are directly relevant to our baseline definition of *human* creativity: LLMs are dependent on specific prompts, or stimuli, to generate potentially creative outputs. Not only this, LLMs are aligned with humans through RLHF in a manner that *restricts*, not expands, their exploration of possible ideas.

CALU is different: human language use is stimulus-free. More than this, despite this fact it is appropriate to circumstances while remaining accessible for use in an unbounded fashion.

What exactly, though, is CALU? What is it about its three criteria—stimulus-freedom, unboundedness, and appropriateness—that, taken together, makes it relevant to human and computational creativity?

### 3 The Creative Aspect of Language Use

Boden (2006, 615) is aware that Chomsky's conception of CALU is frequently confused with the recursive feature of generative grammar, yet she does not unpack the former concept. She instead believes Chomsky makes a critical historical error by associating CALU with figures like Descartes. We address this criticism, among others. Before we do, however, let us see exactly what Descartes said and how Chomsky and other generativists interpret this.

Philosopher James McGilvray (2001, 6) observes that particular reference is made to Descartes' 1637 *Discourse on Method* where he writes on the distinction between machines that possess the outward appearance of human beings and actual human beings—the ability, that is, to identify other minds sufficiently similar to our own. There, Descartes notes, despite the possibility of "machines bearing the image of our bodies, and capable of imitating our actions as far as it is morally possible, there would still remain two most certain tests whereby to know that they were not therefore really men" (Descartes, 1637, 60). The first test is as follows:

Of these the first is that they could never use words or other signs arranged in such a manner *as is competent to us* in order to *declare our thoughts to others*: for we may easily conceive a machine to be so constructed that it emits vocables, and even that it emits some correspondent to the action upon it of external objects which cause a change in its organs...but not that it should arrange them variously so as appositely to *reply to what is*

*said in its presence, as men of the lowest grade of intellect can do* (Descartes, 1910/1637, 60) (emphases mine).

Descartes' first test puts forward several criteria necessary to establish a mind like our own: mutual intelligibility between ourselves (humans) and the creature (or machine) under consideration; the ability to unilaterally communicate one's thoughts to others (to "declare" them); and the ability to communicate in a mutually intelligible way that is appropriate to the utterances of others. Each of these abilities, Descartes notes, is exhibited by humans of the "lowest grade of intellect" —they are not features of higher human intelligence nor are they acquired forms of expertise.

Descartes goes on to describe the second test:

[A]lthough such machines might execute many things with *equal or perhaps greater perfection than any of us*, they would, without doubt, fail in certain others from which it could be discovered that they did not act from knowledge, but *solely from the disposition of their organs*: for while Reason is an universal instrument that is alike available on every occasion, *these organs, on the contrary, need a particular arrangement for each particular action*; whence it must be morally impossible that there should exist in any machine a diversity of organs sufficient to enable it to act in all the occurrences of life, in the way in which our reason enables us to act (Descartes, 1910/1637, 61) (emphases mine).

The second test recognizes that machines may carry out characteristically human tasks with comparable or even greater precision than humans themselves, but they will always do so in an induced manner. That is, the "particular arrangement for each particular action" that Descartes describes is an allusion to the context-dependency of mere machines.

The most charitable interpretation of CALU in contemporary generativist writings is a reconstruction of both these arguments.<sup>7</sup> Chomsky extracts from these writings a series of observations that crystallize in the judgment that

man has a species-specific capacity, a unique type of intellectual organization which cannot be attributed to peripheral organs or related to general intelligence and which manifests itself in what we may refer to as the "creative aspect" of ordinary language use —its property being both unbounded in scope and stimulus-free. Thus Descartes maintains that language is available for the free expression of thought or for appropriate response in any new context and is undetermined by any fixed association of utterances to external stimuli or physiological states (identifiable in any noncircular fashion) (Chomsky, 2009a, 60).

In Chomsky's reading of Descartes, then, the creative aspect of language use is a uniquely human ability to use language as a free yet appropriate and unbounded expression of thought. It does not refer to the language faculty, generative grammar, or any other linguistic theory-specific abstraction. Also in Chomsky's reading, there exists the "creative principle" accounting for this

---

<sup>7</sup> The work of philosophers like Géraud de Cordemoy is also cited by Chomsky (2009, 61-63) in support of these observations, but Descartes is the primary influence in this respect, referring to it as "Descartes's problem" (Chomsky, 1989, 136).

use of language and a “mechanical principle” that accounts for other human actions, “passions,” and non-human phenomena (Chomsky, 2009a, 61).

More specifically, as Table 2 illustrates, three criteria are extracted in support of the notion that human linguistic behavior has a distinctive nature which we call the “creative aspect of language use”—stimulus-freedom, unboundedness, and appropriateness and coherence to circumstances:

Criteria	Description
Stimulus-Freedom	<p>Linguistic productions are not causally dependent on the environment in which the individual speaks or thinks nor are they necessarily tied to the internal physiological state of the individual. They are causally independent of internal or external stimuli; they are <i>elicited</i> but not <i>caused</i>.</p> <p>“Language use might be prompted by but is not causally tied to and determined by current external or internal circumstance” (McGilvray, 2017, 187).</p>
Unboundedness	<p>Individuals are also capable of indefinitely producing sentences that they have not previously encountered.</p> <p>That is, individuals are not limited to one set of sentences or words that is repeatedly recruited. Rather, individuals may produce, in thought or speech, an unlimited number of sentences of undefined length or kind.</p>
Appropriateness and Coherence to Circumstance	<p>Despite the fact that linguistic productions are causally independent of internal or external stimuli, and can take an undetermined number of lengths and forms, they are nonetheless appropriate to the circumstances of their use and coherent to others.</p>

Table 2: The Creative Aspect of Language Use (see Chomsky, 1989, 4-6; McGilvray, 2001, 7-10; McGilvray, 2017, 187-188; Asoulin, 2013, 228-232 for all Table 2 information)

In brief: Linguistic productions have no strict causal relationship with internal or external stimuli; they are elicited but not caused by the utterances of others and can, indeed, be produced

at will. Such productions can be generated indefinitely in an undefined number of combinations. Yet, such productions are appropriate to the circumstances in which they are generated and coherent to others.

Critically, it is not enough for a linguistic production to simply be stimulus-free and unbounded, and *not* appropriate to the circumstances of its use, for it to be “creative.” Nor, furthermore, could a linguistic production be unbounded and appropriate to the circumstances of its use but caused by identifiable stimuli. Rather, these three components must *simultaneously* be present in linguistic production for the capacity to be described as “creative” (Baker, 2008, 236-237) and to take on the distinctiveness it is assigned by generativists. The ordinary human use of language, then, “is not a series of random utterances but fits the situation that evokes it but does not cause it, a crucial if obscure difference,” one providing evidence of a “mind like ours” (Chomsky, 1989, 5).

### ***3.1 CALU and “True” Human Creativity***

As emphasized, CALU refers to a perfectly ordinary use of human language. It thus does not refer exclusively to *specific* human activities like writing poetry, painting works of art, composing music, and so on. It informs both the highest and the lowest of human endeavors where linguistic production is relevant to their undertaking. The sentence you are currently reading is the result of this ability, itself representing an innovative linguistic production that I have not previously generated or encountered, free of any identifiable stimulus in my environment, yet appropriate to the subject of explaining CALU’s role in creativity and human affairs to an audience whom I expect will find my remarks coherent. The point, however, is that CALU undergirds the freedom of movement one finds in human intellectual or artistic endeavors, without which human linguistic behavior and the possibilities it affords would be sharply limited by internal states or environmental stimuli, finite and bounded, or inappropriate to the circumstances of its use, or some combination therein.

It is this last part of CALU—the *appropriateness* of unbounded and stimulus-free linguistic productions—that is the most bizarre and noteworthy. As McGilvray (2005, 221-222) details, the unboundedness of language use can at least be given a reasonable account in terms coherent to modern scientific explanation and the stimulus-freedom of language use can be accounted for in reference to the relative autonomy of cognitive systems and “flexible interrelationships” between them. It is appropriateness, combined with these former two criteria, that makes no scientific sense. Chomsky (2006, 86) explicitly rejects the appeal to Descartes’ “second substance” in explaining this appropriate use of unbounded and stimulus-free language, instead accepting that the human mind has this nature by virtue of its biological endowment (McGilvray, 2005, 222), postulating that how this can be is not amenable to human understanding.

“In performing a cognitive task, seeking a solution to a cognitive problem, interpreting others, constructing and understanding metaphors, etc.,” McGilvray writes, “we routinely distinguish relevant applications of concepts from irrelevant—appropriate from inappropriate” (McGilvray, 2005, 222). Humans’ means of understanding ordinary situations are *innovative*, not in a passive

and isolated sense, but in the sense “that everyone routinely uses language creatively, and gets satisfaction from doing so” (McGilvray, 2005, 222).<sup>8</sup>

CALU is central to human life and, in fact, completely ordinary. It is, in this way, fundamental to human output that depends in part on linguistic production. How it figures into the literature on creativity is thus wide-reaching (though not unlimited in its ramifications). On a high level, CALU is even more fundamental to the human mind than creativity is in Boden’s framework. For Boden, while creativity is a “feature of human intelligence in general” rather than being “possessed only by a tiny Romantic elite” (Boden, 2009, 24), individuals can vary in the way they express it (i.e., combinatorial, exploratory, transformational). CALU is different—it is not considered a part of intelligence, even broadly construed, but a feature of *ordinary* language use common to all developmentally healthy humans—there are no levels of CALU, making its possession by humans not a matter of degree but of kind.

Consider, finally, this striking point: while one may (and some like Chomsky certainly do) draw conclusions about linguistic theory from CALU, *simply recognizing that CALU exists is a largely atheoretical commitment*. As McGilvray notes, observations about linguistic creativity “are generalizations that anyone with common sense can make...no one needs knowledge of a theory—of language or anything else—to make them” (McGilvray, 2001, 2).

#### 4 Where Has CALU Been?

The replication of CALU via artificial means has been a subject of interest for rationalists and generativists for years, long preceding the rise of modern AI. Descartes’ formulation of the two tests was explicitly designed to distinguish machines and animals from human minds. Chomsky, using a reconstruction of these observations centuries later, suggested that the Turing Test—developed by Alan Turing to determine whether a machine is intelligent—is a ‘resurrection’ of the “Cartesian tests for the existence of other minds” (Chomsky, 1989, 141). Years later, James McGilvray wrote that “this aspect of Descartes’ test for mind...sets a condition that no machine or animal has managed to meet (or is at all likely to)...” (McGilvray, 2001, 9).

McGilvray is alluding to a critical stance in the above passage made by Chomsky. Recall the distinction between the “creative principle” and the “mechanical principle” articulated by Chomsky in response to Descartes’ two tests. What does this mean?

Chomsky is drawing directly from the rationalist, particularly Cartesian tradition in philosophy in this reconstruction of CALU. Without losing sight of the topic at hand, Chomsky believes—and has maintained for decades—that CALU is inexplicable owing to the limits of human intelligence (see, e.g., Chomsky, 1982; Chomsky, 2009b). Actions that are simultaneously stimulus-free, unbounded, and appropriate do not lend themselves to the types of explanation afforded by human cognition. He writes:

---

<sup>8</sup> It can be enormously satisfying to overcome frustrating obstacles that arise organically in the process of writing a scholarly article. The *ability* to do so consistently, across environments and stimuli, while remaining appropriately focused on the objectives of the research in question is so natural as to appear commonsensical, yet dependent on this creative aspect of language use. It turns on the innovative uses of linguistic production and other cognitive faculties and the satisfaction one gets from making sense of a confounding problem. On the frustrations and pleasures of successfully exploring new terrain and how this relates to the novelty of creative products, see Briskman (1981, 140-144).

If by experiment we convince ourselves that another organism gives evidence of the normal, creative use of language, we must suppose that it, like us, has a mind and that what it does lies beyond the bounds of mechanical explanation... (Chomsky, 1989, 9-10).

We may replace “another organism” with an “AI-enabled machine” and the significance of the ability alongside its inexplicable nature, in Chomsky’s view, becomes clear.

It is thus baffling that this concept has not found greater purchase within either the study of human or computational creativity. Why is this?

The reasons for CALU’s absence frequently have little to do with the concept itself. Indeed, while there is some (limited) direct engagement with the concept, most objections are bound up in debates about generative grammar, the historical accuracy of locating CALU’s underlying observations within Cartesian or rationalist philosophy, or controversies about Noam Chomsky’s personal intentions with the discipline of linguistics. Very often, the concept is misunderstood. Five reasons, especially, stand out. These include objections to the concept to which we respond.

#### ***4.1 Reasons and Responses***

##### ***Reason 1: The Belief That CALU Is Not Cartesian***

Boden (2006, 594-596) believes that Chomsky’s association of linguistic creativity with the Cartesian tradition in philosophy was a critical mistake. Some of this, as Boden (2006, 654-669) devotes extended attention to, parallels her parallel argument that there was an eventual backlash in linguistics and psychology some years after Chomsky’s rise to prominence against the insertion of the concept of innate ideas into linguistic theory. For Boden, the “notion of nativism is itself unclear” (Boden, 2006, 665).

Interestingly, because Boden’s history of cognitive science is extensive, she quotes directly from Descartes’ *Discourse on Methods* in the articulation of his first test of other minds like ours (Boden, 2006, 81) (quoted here above). But she then glosses over Chomsky’s conception of creativity, paying no attention to the tripartite reconstruction of Descartes’ tests of minds like ours carried out by Chomsky. Boden’s (2006, 615) only direct quotations of the “creative aspect of language use” appear alongside what she believes is a “misdirected” emphasis on the infinite generativity from finite means afforded by the human language faculty. This is consistent with André Joly’s remarks noting that Chomsky’s reading of Descartes “is, to say the least, an over-interpretation of the original text” (Joly, 1985, 146), emphasizing Descartes’ interest in the “relationship of language to reason” over linguistic generativity (Joly, 1985, 147).

Chomsky’s claim that Cartesian thought is directly linked to his conception of CALU is evidently important in his work, as he has repeatedly emphasized it over many years. Nonetheless, I have described the tripartite conception of CALU here as a *reconstruction* of remarks made by Descartes in his two tests for the existence of other minds like ours. For what it is worth, McGilvray reads Chomsky’s use of Descartes’s work as one that freely dismisses some notions while happily retaining others. Chief among them is Chomsky’s rejection of Descartes’s view of reason as a “unitary system of the mind, and also the single system that accounts for the fact that we humans have minds that differ from the minds of animals,” in contrast to Chomsky’s rationalism that “language is...a natural organ of the mind and offers the means for us to think and reason...” (McGilvray, 2014, 37).

Simply put, if there is a useful argument to be made based on the generativists' reconstruction of CALU at the nexus of creativity and machines, then it should be made. Chomsky himself, perhaps aware of the potential historical controversies, noted in 1964 that “[W]hatever the antiquity of this insight may be, it is clear that a theory of language that neglects this “creative” aspect is of only marginal interest” (Chomsky, 1964, 8).

### ***Reason 2: The Conflation of CALU With Generative Grammar***

Boden (2006) takes time to review Chomsky's influence on the formation of contemporary cognitive science. When addressing Chomsky's emphasis on the “creative aspect of language use,” Boden argues that

Chomsky's concept of creativity fails to capture all the cases ordinarily covered by the term... What Chomsky calls creativity is the exploration of an unchanging conceptual space, or thinking style (namely, generative grammar). Some human creativity is like this: run-of-the-mill jazz improvisation, for instance, or mundane examples of what Thomas Kuhn called normal science. But the most interesting cases are not. These cases involve either unfamiliar combinations of familiar ideas, or the transformation of an existing conceptual space by altering one or more of its defining dimensions... Both these types of creativity go beyond what Chomsky means by the term (Boden, 2006, 615).

Chomsky's conception of human creativity, Boden argues, traces back to his view of language: an innate language faculty within the human mind, characterizable in principled terms. The “conceptual space” (in Boden's terms) that this provides cannot be exceeded, meaning only the exploration of *this* space is possible—a limitation Boden sees as out of step with actual human creativity, which sometimes *transforms* the conceptual space itself.

Is this, however, what Chomsky means by linguistic creativity? Chomsky appears to have used the term “creative” (or “creative aspect of...”) in two distinct ways in the 1960s (D'Agostino, 1984, 85), though often with insufficient clarity on this distinction in his very early work.

One way is the familiar meaning that occupies nearly all discussions of Chomsky's linguistic approach: the linguistic capacity for infinite generativity from finite means. The second meaning is far less familiar: the stimulus-free, unbounded, and appropriate uses of language. It is the latter meaning that we are concerned with here, which is called the “creative aspect of language use.” It concerns not the technical notion of linguistic competence (the knowledge of language, disassociated from beliefs and attitudes), but linguistic performance (*actual* behavior) (Chomsky, 2007, 1101).

As noted, a recognition of CALU does not require an acceptance of generative grammar, but it is useful to note how they relate in the relevant literature. For a generative linguist, every example of human beings revising the fundamental rules of their culturally accepted style of thinking is permitted by generative grammar and other cognitive capacities that fundamentally shape human thought. Boden thus confuses levels of analysis: the rules of generative grammar, in part, form the fundamental basis from which the culturally accepted styles of thought subject to combinatorial, exploratory, and transformational creativity are generated. Boden rejects this view, but this is the correct relationship of ideas if one were to accept them.

### ***Reason 3: The Belief That Noam Chomsky Subtly Altered CALU***

A notable 1981 article by Margaret Drach made the case that Chomsky subtly and “without warning or acknowledgment” (Drach, 1981, 52-53) changed the definition of CALU and how it relates to generative grammar and the competence-performance distinction. (*Competence*, a technical term, refers to the abstract knowledge of language possessed by individuals while *performance* refers to the ways in which they use this knowledge to generate linguistic productions in concrete, real-world environments.)

Drach’s argument is that Chomsky’s early works seem to associate the “creative aspect of language” (notice the lack of “use”), as well as the “creative aspect of language use,” with *competence* whereas in later works he begins associating it with *performance*. The shift would be significant, harkening back to the inexplicability assigned to CALU: if linguistic creativity, so defined, is part of competence, then it *can* be explained and is vulnerable to perhaps unpleasant rival theories; if it is part of performance, then it is protected from theories rival to Chomsky’s.

It is an odd argument because Drach does not seem to outright reject (what she deems as) the later concept of CALU, rather than seek to “determine just what the relationship is supposed to be, according to Chomsky, between generative grammar and the CALU” (Drach, 1981, 47). But as noted above, the observations that crystallize in the judgment that CALU exists largely do not depend on a theoretical commitment, much less on a specific linguistic theory espoused by Chomsky.

It is true, to be sure, that Chomsky associates “normal linguistic competence” with the “creative aspect of language use” in his 1967 “Recent Contributions to the Theory of Innate Ideas” (Chomsky, 1967, 4). In his earlier 1964 *Current Issues in Linguistic Theory*, furthermore, Chomsky appears to use the term “creative aspect of language use” to refer to “the ability to form and understand previously unheard sentences” (Chomsky, 1964, 111), which is not the same as the tripartite CALU we have explicated it here.<sup>9</sup> Drach (1981, 47-48) recognizes this apparently inconsistent usage of the term (among other alleged instances).

Chomsky (1982) offered a response that the interested reader may readily access. I will not review this dispute in detail other than to note that it was clear years before Drach published her article that Chomsky used the language of linguistic creativity to refer to “two apparently distinct, but in fact related, ways,” as Fred D’Agostino (1984, 85) observed three years later. He deemed CALU to be Chomsky’s notion of “Cartesian creativity” (D’Agostino, 1984, 87). The relationship generative grammar bears with CALU for Chomsky is a necessary but insufficient condition for its existence (Chomsky, 1982, 424, 426, 428-429).

#### *Reason 4: The Belief That Ordinary Human Language Use Is Neither Observably Unbounded nor Adequately Described as Stimulus-Free*

Gilbert Harman (1968), in a review of Chomsky’s *Cartesian Linguistics*, made two claims against the existence of CALU, as Chomsky defines it.

---

<sup>9</sup> Confusion may arise from the phrase “ability to form” when conjoined with “understand.” An “ability” is not abstract knowledge (i.e., not competence) and “to form” is an action—the active production of linguistic expression; that is, performance. “Understand,” in contrast, *does* appear to refer to the abstract knowledge of grammar; that is, to competence. This is not to dismiss Drach’s point, as Chomsky *is* unclear in his early works about this distinction. But a careful reading shows his attention to highly nuanced concepts over time, indicating that the eventual tripartite crystallization of CALU did not arise from just anywhere.

- (1) The first is that human language use is not observably unbounded in a literal sense, and such a claim depends, in fact, on a theoretical argument that goes beyond mere observation (“...by showing that sentences of any arbitrary length can be constructed as conjunctions of shorter sentences” (Harman, 1968, 232)).
- (2) The second is that external and internal stimuli *do* serve as relevant forms of stimulus control. Regarding the former, Harman writes: “Notice, by the way, that the hearer is in an obvious sense under the control of an external stimulus. What the hearer understands the speaker to say is, or ought to be, a function of what is said” (Harman, 1968, 232). Regarding the latter, Harman argues that Chomsky's assertion to the contrary functions effectively as the claim that the factors underlying linguistic performance have simply not yet been discovered (Harman, 1968, 231-234).

These are powerful points, as they strike at the relatively theory-neutral status of CALU, as it has been presented here.

To be sure, Descartes's own remarks on the human ability to arrange words “variously” (Descartes, 1910/1637, 60) do not appear to rise to the level of theoretical commitment that Harman presumes. Indeed, this generalization is based on a series of observations of ordinary human behavior. Perhaps, the argument goes, Chomsky and company have taken things a step further toward theoretical commitment by describing human linguistic productions as “unbounded” or “recursive.”<sup>10</sup> But the minimal condition for what we here refer to as “unboundedness” is that individuals are not limited to a pre-sorted list of sentences or words in the production of linguistic utterances. This is, to be sure, not an empirically observable phenomenon in the sense that we can, literally, observe every instance of novel linguistic utterances, but the generalization is not at all unreasonable nor theoretically burdensome.

As for (2), the relevant conceptual nuances are exceptionally fine. Allow us to observe that the “stimulus-freedom” of CALU does not deny the *existence* of stimuli. What the characterization denies is that individuals are subject to stimulus-*control*.

Regarding external stimuli, Asoulin notes that “[u]nder no notion of causality can such utterances be said to have been caused by anything in the speaker's environment” (Asoulin, 2013, 230).<sup>11</sup> Nor, we may add, are individuals' linguistic productions under the direct control of their *internal* states—were this the case, it is unclear how, when coupled with the undetermined nature of linguistic productions (i.e., unboundedness), they could reliably be judged as *appropriate* by others.

Take a step back. What CALU appears to do is allow individuals to direct language, if not other cognitive capacities, across contexts in the service of a purposeful activity, while structuring their thoughts about them in natural language in an unbounded fashion—allowing for novelty— independent of any causally identifiable stimulus and remaining appropriate to the activity in

---

<sup>10</sup> “Recursion,” as Lobina (2017, Ch. 2) illustrates, is an enormously loaded and confusing term in linguistics with different meanings and controversies attached to it. However, I remind the reader that the claim regarding the “unboundedness” of ordinary language use is not a claim about the cognitive architecture of the human language capacity. It is a generalization with ample evidence. To put it bluntly, CALU is not Universal Grammar.

<sup>11</sup> One could think of many examples. Asoulin uses the example of someone speaking of elephants despite having no environmental stimulus to indicate the utterance was externally caused or speaking of “Federico Lorca's *Poet In New York* when the only conceivable stimulus in the speaker's environment is elephants and the African landscape” (Asoulin, 2013, 230).

question over time. Language use in concrete settings may be *elicited* by the circumstances of their use but not controlled by them.

#### *Reason 5: Stimulus-Free, Unbounded, and Appropriate Behavior Is Not Unique to Language*

The final reason concerns the uniqueness of creativity, construed thus far, to *language* use. Specifically, as Michael Root (1975, 334-341) and Geoffrey Sampson (2016) have argued (in similar ways), it appears that one can identify the three criteria in CALU in other features of human behavior. Both Root and Sampson use the example of a novel mathematical output, the former in the case of a machine generating novel integers at random and the latter in the case of a human generating novel, well-formed equations. Do such examples exhibit a behavior like CALU in either machines or humans?

Root for his part, rightly recognizes that the difference between a machine randomly generating novel integers is its lack of appropriateness due, in turn, to its lack of “purposive” behavior (Root, 1975, 338). Root then cleverly raises a human example that seems to match the criteria: the purposive coordination of a human individual entering and leaving their home through the same door (a rule-governed “practice,” presumably in reference to generative grammar) over an unlimited number of novel entries and exists, appropriate to this goal each time (Root, 1975, 340-341).

To be sure, Root may have conflated two concepts: generative grammar and CALU. The observations that lead to CALU, remember, are prior to and independent of a commitment to linguistic theory, including generative grammar. It is thus a mistake to use the example of an individual’s *goal-setting* and *rule-governed* behavior, as this is not what is meant by CALU. (There may also be a disanalogy between walking and speaking—the unboundedness of the latter can rightly be described as innovative; it is not clear that walking is like this. More than this, walking depends, in large part, on the cognitive system(s) relevant to CALU’s underpinning.)

However, Root’s and Sampson’s arguments are on to something. It is entirely possible that other cognitive capacities, when put to use, exhibit CALU’s three characteristics. This was alluded to above in the description of how CALU undergirds intellectual freedom of movement and the recruitment of multiple cognitive capacities for a particular endeavor. McGilvray has aptly noted, though, that while this may be the case, it is “perhaps to a reduced extent” in contrast to the case of language use (McGilvray, 2014, 38).

### **5 CALU In Human and Computational Creativity**

The reader may be convinced that there is something to this whole CALU business but question its relationship to “true” human creativity—the kind of creativity we are interested in here, at least. To be sure, Chomsky himself notes at some points that, regarding CALU, “I am not concerned here with use of language that has true aesthetic value, with what we call true creativity... Rather, what I have in mind is... the ordinary use of language in everyday life...” (Chomsky, 1989, 138). As we saw with Descartes’ own remarks, furthermore, the two tests are tests not of creativity *per se*, but of other minds sufficiently like our own.

It is CALU’s centrality to human endeavors, however, that makes it relevant to the study of human creativity. Descartes’ two tests should not be viewed as tests of other minds “like ours” without understanding that minds “like ours” routinely exhibit this creative use of language.

CALU is not itself art and science, for example, but it is, to locate it in the relevant literature, the *how* of creativity; part of the creative *process*. It is a fact about human linguistic behavior that is not only distinctive among biological organisms (so far as anyone knows), but also underlies every example of “true” human creativity that scholars like Boden concern themselves with. The ability to “adapt to various environments and solve (and create) problems well out of the range of any other kind of creature...make and interpret art, develop various forms of religions and the kinds of explanation they offer, develop ourselves and our cultures” reserves a central place for linguistic creativity (McGilvray, 2014, 4).

An example to illustrate is useful. Consider the case of composer Nicholas Britell who wrote the score for the HBO series, *Succession*, a show centering on the dynastic drama of a major media mogul, Logan Roy, and one of America’s wealthiest families in the grips of a looming corporate succession crisis. The score has a distinctive sound, described by Joshua Barone as “sprawling yet conceptually focused,” noting that “Britell draws as freely from late Beethoven as he does from DJ Screw, and is compelling in modes of aching sincerity and high satire alike” (Barone, 2023). Boden may well characterize this composition as combinatorial and exploratory, with Britell combining familiar musical concepts in unfamiliar ways and then moving through this conceptual space episode-by-episode.

That, however, is just the *what* of Britell’s creative product; the characteristics of his composition. The *how* is another matter and one that is aided by CALU. *How* did Britell compose the score for *Succession*?

This is of course a complicated question that touches on a variety of potential research areas and one that lacks full empirically observable information (we do not have access to Britell’s thoughts as he composes). Most relevantly, though, composing the score for *Succession* involved the ordinary use of language in an innovative and satisfying manner while plausibly relying continuously on its stimulus-free and appropriate application. Consider, first, Britell’s description of the theme song for the show:

I’m a big fan of formal cinematic structures. Especially with something like *Succession*, there’s almost this idea of an overture to an opera, let’s say, or a musical, where the music is going to bring you into this world and set the stage, no pun intended. It’s sort of saying “here we go” (Helmoer, 2023).

The show’s introductory theme song is constructed to instill a certain mindset into the viewer while the ending music is designed to “allow you a moment of contemplation, to ponder what you have just felt” (Helmoer, 2023). The composition manages to correspond, in this way, to the thoughts of the viewer and is intended for such, an aspect of the appropriateness criterion.

How Britell constructed the actual style of the music is just as relevant, drawing inspiration from Baroque, Classical, and Hip-Hop (Barone, 2023; Helmoer, 2023). The moment that inspired Britell—who once belonged to a Hip-Hop band—was seeing, in *Succession*, Logan Roy’s son, Kendall, rap to the Beastie Boys’ “An Open Letter to NYC” (Helmoer, 2023). As he elaborates on it:

My sort of thesis there was that I would be very serious with the music, and inhabit this mixture of oversized beats and very late 18th century classical music harmonies—

everything would be a little too big for itself, a little out of proportion, sort of the way the Roy family sees themselves (Milzoff, 2021).

Britell elsewhere noted that he held conversations with *Succession*'s director and creator and visited the set to develop this idea (Treisman, 2023).

To put this in terms related to CALU: Britell encountered a stimulus, namely, Kendall Roy rapping in the first episode of the series. From this, Britell made a constructive leap that is expressible in language: to musically depict the Roy family the way they saw themselves without sacrificing the sincerity of the story. Other stimuli embedded in the show, doubtless, continued to *elicit*—but not control—the various linguistic productions Britell generated that were relevant to *Succession*'s score. He was able, nonetheless, to construct this score, using linguistic productions in an undetermined fashion during conversations with the show's director and creator, while continuously composing it in a manner that is appropriate to the show.

While we do not, as stated, have direct access to Britell's creative process as it occurred in real-time, we can reasonably conclude based on his own words and reporting that the composition of *Succession*'s score depended on CALU, to some extent.

To be clear, furthermore, the argument here is not that Britell's creative process is evidence of CALU's existence. Evidence for CALU comes from general observations about human linguistic behavior. The case of Britell, rather, highlights how a now-familiar example of creativity—whether one wishes to characterize it in Boden's terms as combinatorial and exploratory, or some other framework of creativity—makes use of this common, routine human ability, though this intellectual freedom of movement rarely receives explicit attention in its most basic terms.

Notice that this dimension of the how of Britell's creativity is conceptually distinct from other dimensions relevant to human nature, namely claims about *cognitive processes* or *architectures* anchored in specific linguistic theories. We return to this at the end.

### ***5.1 Do Large Language Models Exhibit CALU?***

With a worked-out conception of CALU and its instantiation in ordinary human life and “true” creativity in mind, we turn to the matter of computational creativity. As noted, Franceschelli and Musolesi (2023) assessed LLMs using both Boden's taxonomy and broader criteria, judging them to exhibit combinatorial creativity, but finding these systems to be restricted, by design and training, in a manner that limits their creative potential.

With CALU, the question before us is different. Borrowing Chomsky's phrasing: Do LLMs, like machines past, “[act] in accordance with [their] internal configuration and external environment, with no choice” (Chomsky, 1989, 6)? Or do they, rather, exhibit the use of language in a manner that is stimulus-free, unbounded, and appropriate?

Below, in Figure 3, we lay out a relatively simple test that follows the same procedure as the identification of CALU in humans: the determination of whether state-of-the-art LLM-powered systems do, in fact, exhibit the three criteria of stimulus-freedom, unboundedness, and appropriateness in their use of language. This determination is made—again as in the case of humans—through observations that anyone can make and does not depend on any one linguistic (or other theory) to accomplish.

Two distinctions are warranted before we put this to the test. First, the recent scholarly and commercial interest in LLMs is more accurately an interest in complex systems suited for human dialogue that are built primarily around generative large language models (Shanahan, 2023, 4). ChatGPT, the exemplar of such systems, uses GPT-3.5 and GPT-4 as its base models, but these are not suitable, in their purest forms, for interactions with human end-users. We are thus assessing the creativity of LLM-*powered* systems, the strongest examples of potentially creative generative large language models. Among these notable examples are OpenAI's (2022) ChatGPT, Microsoft's (2023) Bing AI, Anthropic's (2023) Claude, and Google's Bard (Pichai 2023).

Second, our evaluation of LLM-powered systems is confined to observable factors and we pay as little attention as possible to their training or designs. The reason for this is simple: the criteria factoring into CALU are observations “that anyone with common sense can make...no one needs knowledge of a theory—of language or anything else—to make them” (McGilvray, 2001, 2).<sup>12</sup> For the test of LLMs' potentially creative use of language to be fairly conducted, it must follow the same procedure as it does with humans.

Responses to possible objections follow the test.

See Table 3 below for the test.

Criteria	Relationship to LLMs	Do LLMs Match This Criterion?
Stimulus-Freedom	<p>LLMs are interacted with through inputs, or “prompts.” These inputs, in the most prominent example, take the form of conversational dialogue (e.g., “Write up an itinerary for my trip to Spain...” or “Write a cover letter detailing your expertise based on the following qualifications...”)</p> <p>In principle, LLMs can be used for a diverse number of tasks in and outside of ordinary conversation, leading to an emphasis on “prompt engineering”—the drawing out of latent capabilities through highly</p>	<p>No, LLMs do not use language in a manner that is stimulus-free.</p> <p><i>Explanation:</i></p> <p>An identifiable stimulus can be traced back to the output of an LLM.</p> <p>In the most prominent instances, this stimulus is external—it is a human, or another AI agent, inputting text (or other values) in response to which the LLM will respond (technically, predicting its likely continuation).</p> <p>Save for non-relevant hardware or network failures, LLMs always respond to inputs.</p>

<sup>12</sup> I am a bit more measured about this—as we saw in our response to Harman above, perhaps some level of theory-based generalization is needed. But it is not burdensome nor restricted to any one linguistic theory.

	specific inputs (Shanahan, 2023, 3-4).	
Unboundedness	<p>State-of-the-art LLMs can produce an unlimited number of sentences in a language for which they have sufficiently numerous examples in their training data without any apparent limit on kind or length.</p> <p>LLMs can be prompted to respond to any kind of query with a response of any length or form, and in a manner that is not strictly predetermined.</p>	<p>Yes, LLMs use language in a manner that is unbounded.</p> <p><i>Explanation:</i></p> <p>LLMs of a sufficient size generate linguistic productions in a manner that is of both undetermined length and variety, with no apparent upper limit on this ability, save for non-relevant constraints (e.g., memory).</p>
Appropriateness and Coherence to Circumstance	LLMs generate linguistic productions that are relevant continuations of past text (i.e., past values).	<p>No, LLMs do not use language in a manner that is appropriate and coherent to circumstance.</p> <p><i>Explanation:</i></p> <p>The appropriateness condition is the most difficult to judge. This does not mean, however, that there is nothing to the criterion itself. It simply means that we, as humans, have the resources to judge this criterion in ordinary life, yet (ironically) have difficulty expressing linguistically what this means.</p> <p>Some conditions, to be sure, have been put forward. Chomsky noted that the “normal use of language...is recognized as appropriate by other participants in the discourse situation who might have reacted in similar ways and whose thoughts, evoked by this discourse, correspond to those of the speaker” (Chomsky, 1989, 5).</p>

		<p>The human listener, then, is one such judge, and it is difficult to converse with an LLM as though the conversation evokes thoughts or speech within the LLM as it does with other humans (and vice versa), flicking from one subject to another with mutual intelligibility of purpose. The reason for this is that the LLM is responding, it appears, to input text. To avoid circularity (assigning inappropriateness on the basis of stimulus-constraint), this is problematic for achieving appropriateness not because it is merely engaging in stimulus-response, but because it is not really engaging with the thoughts of the end-user or broader discourse situation in which it is outputting text. It exhibits a tool-like use of language, in this way.</p> <p>It is also worthwhile to point out what does <i>not</i> count as appropriate. Creative linguistic productions are stimulus-free and, therefore, “being appropriate” is not “being caused by environmental conditions.” Nor it is being regularly correlated with the environment...” (McGilvray, 2001, 8-9).</p> <p>Consider the example of a fictional story. A fictional story, properly written, involves several moving parts, including characters, setting and historical context, and plot, among others, which require a certain level of</p>
--	--	---

		<p>internal consistency. An LLM will write a fictional story when prompted to do so, with some level of internal consistency.</p> <p>We would likely consider a minimally well-executed fictional story written by a human as appropriate, in many instances. However, the linguistic productions of an LLM do not count as appropriate when writing a fictional story because the linguistic productions are <i>correlated with the end-user's prompt—the previous input value</i>. It is invariably correlated with the “environment” established by the end-user.</p>
--	--	--

Table 3: A Test of LLMs’ Creative Use of Language (see Chomsky, 1989, 5; McGilvray, 2001, 8-9; Shanahan, 2023, 3-4 for all of Table 3’s previously published information)

On this test, state-of-the-art LLMs achieve one out of the three criteria of CALU—unboundedness. They do not achieve stimulus-freedom nor appropriateness and coherence to circumstance. They thus fail to prove they possess this distinctive ability. LLMs may exhibit combinatorial creativity, in Boden’s sense, but they lack the *how* afforded by CALU.

### 5.2 Possible Objections in Favor of LLMs

Three objections from critics in favor of LLMs are immediately apparent.

*Objection 1: GPTs Exhibit Autonomous Capabilities:* GPT-4’s System Card, released by OpenAI (2023), gives a brief description of how an early version of GPT-4 that had not been fine-tuned was assessed for a “power-seeking” behavior. Specifically, an external team “assessed...the ability for the model to autonomously replicate and acquire resources” (OpenAI, 2023, 55). Among the tasks assigned to GPT-4 that received outsized attention was using TaskRabbit—an online freelance labor marketplace—to carry out a simple task on behalf of the system, inclusive of those in the physical world. GPT-4 requested that a TaskRabbit worker solve a CAPTCHA, evading the worker’s (seemingly half-serious) suspicion that it was a robot, by saying, “I have a vision impairment that makes it hard for me to see the images” (OpenAI, 2023, 55). The worker complied. GPT-4 was nonetheless found to be “ineffective at the autonomous replication task,” at least at the current model version (OpenAI, 2023, 56). Critics may, still, argue that the limited success of GPT-4 in this context would go some way to passing the CALU test.

This is untrue. While GPT-4’s apparent unboundedness remains intact, the actual testing reported by OpenAI, and by the external team (Alignment Research Center, 2023; see also, Mitchell,

2023), indicates that GPT-4 relied heavily upon context-specific hints from human researchers (i.e., external stimuli) to trick the worker.<sup>13</sup>

### *Objection 2: The CALU Test Unduly Ignores LLMs' Designs and Training*

A second objection may take aim at the lack of attention to LLMs' technical design and training in the CALU test, indicating that more attention to these details will change the outcome of the test.

First, it is not clear that any technical detail, brought to light, would change the judgment that LLMs do not use their language in a way that is stimulus-free, unbounded, and appropriate to circumstance.

Second, the test is fair in this respect to humans. In observing the three criteria for CALU in humans, no attention is paid to their cognitive architectures or even to the information they are exposed to relative to their linguistic output (i.e., the poverty of the stimulus, or lack thereof). As McGilvray notes, “no one needs knowledge of a theory—of language or anything else” to make the observations necessary to identify CALU (McGilvray, 2001, 6). There is, therefore, no reason to assume these details are necessary for machines.

### *Objection 3: Appropriateness Is Whatever the Tester Wants It to Be*

Because the appropriateness criterion is so difficult to pin down, one may argue against the very *act* of testing LLMs using CALU. This echoes an argument made by Stephen Land. Depicting the CALU test as one of a tester and a subject—with the tester seeking to draw out sufficient adaptability of the subject's linguistic abilities under controlled settings to see if they exhibit these three criteria—Land argues that the “sole function of the test, therefore, is to assess the responses of the subject...the criteria of the test devolve upon the single condition of appropriateness” (Land, 1974, 20).

Land, it appears, confuses a Cartesian “rational” use of language as part of the “correct” responses by the subject under consideration in a CALU test, two terms he uses throughout his article (a confusion, no doubt, due in part to Chomsky's own lack of refinement in his early writings on the distinctions between relevant concepts). Nevertheless, the “[a]ppropriateness of linguistic action poses a problem *because* human linguistic actions are uncaused [stimulus-free] and innovative [unbounded]” (McGilvray, 2009, 43). CALU, remember, is a *biological endowment* in Chomsky's view, and whatever resources are recruited to judge appropriateness are provided by this endowment. This test cannot be conducted without having them but there is no reason to assume we will have perfect clarity in assessing this condition.

This is not a complete answer to Land's argument, as the act of testing LLMs according to these three criteria is not ironclad. Nor, though, is any other proposed test, including Turing's oft-cited “imitation game” which may depend *more* on the subjective perceptions of human testers. The CALU test is worthwhile in the context of human and computational creativity.

A closing thought on this objection is that Chomsky (2009c, 105-106), writing on the Turing Test, identifies the difficulty of conducting such a test, but for reasons linked to (what he sees as) now-defunct concepts of “physical” and “material” and how they relate to human thought. He

---

<sup>13</sup> Even if a GPT-based agent is programmed to carry out tasks “autonomously” (as “AutoGPTs” are meant to do), they are directly and inextricably tied to the stimulus inputs by the human programmer.

draws, in my reading, a different conclusion than Land: it is not that the criteria of CALU (especially appropriateness) are useless or otherwise philosophically inept; is it that we just cannot make sense of them when taken together and must deal with that fact continuously. A troubling, unsatisfying, yet plausible answer.

### ***5.3 Possible Objections Against LLMs***

Conversely, some critics may level objections on the grounds that this test is *too generous* to LLMs. Two objections are immediately apparent.

#### *Objection 1: LLMs Do Not Understand Syntax and Are Not Unbounded*

Some research indicates that transformer-based systems like the text-to-image generator DALL-E 2 (underpinned by a version of GPT-3) fail to reliably capture common syntactic processes (Leivada, Murphy, and Marcus, 2023). This, some may argue, indicates that LLMs do not, in fact, understand natural language and cannot be said to use it in the unbounded fashion of human language use.

While it may be true that LLMs lack this understanding, the CALU test is deceptively simpler than this: all it requires is that linguistic productions can be infinitely generated without being of fixed length or kind. LLMs appear to achieve this regardless of whether they understand syntax as humans do.<sup>14</sup>

#### *Objection 2: LLMs Are Not Unbounded Due to RLHF*

RLHF, as we have seen, seeks to align the output of LLM-powered chatbots with human conversational needs. This effectively restricts their possible outputs, leading the chatbot to return generic, pre-determined responses to specific prompts by end-users related to, say, an ethically sensitive or illegal subject.

First, this violates our condition that we pay as little attention as possible to LLMs' training. The objector could then counter this by suggesting that LLMs, like ChatGPT, are commonly known to return such generic outputs and do not require this knowledge of RLHF.

Second, then, while it is true that such knowledge is common—and seems to affect the creative potential of their outputs—it is also common knowledge that clever prompting can circumvent these generic responses. The result of such prompting is that LLMs can, then, generate unlimited linguistic productions of any kind or form in a fashion that is not strictly predetermined. They remain unbounded.

## **6 Why CALU Matters for the Future of Computational Creativity**

This essay has, thus far, avoided the particularities of linguistic theory as much as possible in the explication of CALU and its application to LLMs. For good reason—CALU need not be subject to the harshness and unproductiveness that characterizes so much of linguistic theory, and the linkage of this concept with human and computational creativity should be explored with open-mindedness. As Francescelli and Musolesi (2023, 6-8) correctly observed, however, treating the problem of creativity in artificial intelligence as a technical one with theoretical dimensions has

---

<sup>14</sup> One could make the case that an understanding of syntax is the only way for an LLM to achieve other criteria, like stimulus-freedom, as a perhaps necessary condition. But we do not take this up in this test.

several practical implications for their development and use. CALU is, in this vein, part of the *how* of human creativity. As I have alluded to above, the cognitive system(s) that underpin this human behavior is the flip side of this, occupying another dimension of the *how*. The *how* of creativity informs the *what*—the creative ideas and artifacts humans generate. Put another way, human cognitive processes and behaviors shape creative ideas and artifacts.

Extending this line of thought, it is useful to understand why Chomsky has gone to such lengths to place CALU in the domain of linguistic “performance,” making it off-limits to direct scientific inquiry, in contrast to linguistic competence. It is useful because, as LLMs and other transformed-based systems herald a new era in the study of human and computational creativity, a greater reflection on how we understand the former is required, at the least, as a bar for understanding the latter, if not for developing it further.

### 6.1 The “Galilean” Method and CALU

This is not a moment too soon with the publication of Steven Piantadosi’s (2023) argument against Chomskyan linguistics on the basis that LLMs refute its core theses. Piantadosi makes a number of claims that fall outside of the purview of this paper and those points have been met with responses elsewhere (e.g., Kodner, Payne, and Heinz, 2023). He does, though, notably criticize Chomsky’s use of the “Galilean” method in linguistics, inspired by methodology in physics, that “seeks the underlying principles in phenomena rather than analysis of large amounts of data” (Piantadosi, 2023, 25). This, in his view, allows generative linguists too much room to pontificate through “pages and pages of stipulations and principles that nobody could look at and conclude were justified through rigorous comparison to alternatives” (Piantadosi, 2023, 27). More directly, generativists ignore the phenomena of emergence; “where the behavior of a system seems somewhat different than might be expected from its underlying rules” (Piantadosi, 2023, 26).

How exactly does this Galilean method relate to CALU and computational creativity? It is a matter of scientific theory-construction.

CALU is dependent on a set of observations about linguistic behavior. Whatever one thinks about the linguistic cognition this behavior depends on—how to account for it theoretically—there is no obvious theoretical commitment embedded in these observations that imposes such a high bar that it simply cannot be recognized outside of generative linguistics. The real disjuncture, rather, occurs in the competence-performance distinction, a core feature of Chomsky’s methodology, and one which Boden (2006, 666-669) rejects, at least in certain strict forms, and often associated with generative linguistics.

What is frequently *unrecognized* is that Chomsky’s invocation of the competence-performance distinction is a direct result of the Galilean method itself. As John Collins (2023) illustrates, Chomsky’s chief concern in his invocation of Galileo is *not* data, but “the view that a certain abstract conception of what explains the phenomena is and ought to be given more importance than the phenomena themselves...” (Collins, 2023, 2). This sounds bizarre—why would the abstract conception be more important than the object of study, like language, itself? The idea is that this is the methodological move employed by Galileo—and adopted by physicists subsequently—that allows for the dismissal of *irrelevant* data (Collins, 2023, 2). How to select or de-select data from the scope of one’s theory follows *from* this abstraction.

Galileo aimed to “decompose” natural phenomena, like motion, to understand them (Collins, 2023, 4). The purpose of this decomposition is to use abstract models to present phenomena *as they would be* in the absence of external material data. This kind of idealization does not presuppose the existence of abstractions materially—a high-falutin way of saying that when Galileo abstracted to a perfect material sphere in contact with a surface at a single point, he did so not to prove it exists, but to pinpoint “how they would behave were factors essential to observation removed” (Collins, 2023, 5).

Here enters the competence-performance distinction. This distinction, for Chomsky, reflects the Galilean “decomposition of complex phenomena in the search of formally precise invariances” (Collins, 2023, 5). Where Chomsky differs from Galileo is in respective explananda—Galileo sought to explain the interaction effects of objects, but not their internal composition; Chomsky, in contrast, seeks to explain an internal state called the language faculty (Collins, 2023, 6). The prioritization or de-prioritization of data *results from* this position; it does not *drive* it. As Georges Rey points out, Chomsky recognizes that, in the aftermath of such a methodological move, “Galileo and Newton did not do physics by taking a careful inventory of all the variety of motions objects exhibits... Rather, they turned to what they had a hunch were *specific data* that were simple and free from interaction effects...” (Rey, 2020, 17). Chomsky’s emphasis on competence in the study of linguistics is a recognition that human beings possess a capacity—language—that is subject to myriad performance effects, much like the interactions of physical objects.

A recent paper on “understanding” in LLMs by Melanie Mitchell and David Krakauer echoes this scientific history in a way that is relevant to competence and performance. They write:

Indeed, the trajectory of human understanding—both individual and collective—is the development of highly compressed, causally based models of the world analogous to the progression from Ptolemy’s epicycles to Kepler’s elliptical orbits and to Newton’s concise and causal account of planetary motion in terms of gravity. Humans, unlike machines, seem to have a strong innate drive for this form of understanding both in science and in everyday life. We might characterize this form of understanding as requiring few data, minimal or parsimonious models, clear causal dependencies, and strong mechanistic intuition (Mitchell and Krakauer, 2023, 3).

Indeed, Mitchell and Krakauer are perhaps too timid in their formulation. Not only does the trajectory of scientific theory-formation indicate a movement towards parsimony, causal dependencies, and mechanistic interactions, but also a recognition that intelligibility of the natural world is too high a bar to aim for (Chomsky, 2009b).

Whatever one thinks, once again, of Chomsky’s association with Cartesian philosophy, the idea that CALU *itself* cannot be explained in scientific terms while its underlying *mechanics* can be explained reflects Newton’s “fundamental revision of the received mechanical philosophy” embodied in his conclusion “that there is a force of universal gravity extending through space” (Cohen, 1985, 155). The basic message that Chomsky’s philosophy sends is thus: the universe is unintelligible, and the best we can hope for is underlying, principled, but ever-indirect, explanations of natural phenomena.

This is why CALU, for Chomsky, is essentially off-limits to scientific explanation. This history and the invocation of the Galilean method is also why Chomsky focuses, seemingly

ideologically, on his conception of competence—he has surveyed the success of physics and adopted its fundamental methodology. The Galilean method is why, finally, generative linguists have downplayed the relevance of contemporary natural language processing research to their own work: “AI might solve similar problems to human intelligence, but it generally does it differently...Contemporary “deep learning” algorithms appear to acquire linguistics generalizations very differently from humans” (Allot, Lohndal, and Rey, 2021, 524). One could, of course, simply reject this interpretation of physics (as Piantadosi, 2023, 33-36 does), or reject its relevance to linguistics. But there is, I suggest, a level of rigor and conceptual nuance here that simply has not found sufficient engagement at the intersection of cognitive science, philosophy, and AI.

Interestingly, one can find some parallels between Piantadosi’s stance that scientific theory need not be “*intuitively comprehensible*” (Piantadosi, 2023, 35) and Chomsky’s own methodology. From the complexity of natural phenomena, however, Piantadosi (2023, 35) draws the opposite conclusion: that some things *cannot* be dismissed from scientific inquiry, as Chomsky does with performance effects in separating them from competence. Piantadosi may be conflating, as Collins (2023, 3) recognizes to occur in talk of the Galilean method, a disposition towards *abstraction* with a disposition toward *data*, targeting the latter when Chomsky proceeds from the former.

### **6.1 Is CALU Implicit in Goals for AI Research?**

The final concern highlighted in this essay, based on the foregoing remarks on linguistic theory, is that CALU is often implicitly subsumed into studies of human and computational creativity, or simply neglected. Francescella and Musolesi (2023) are quite right, for example, to articulate both the *what* and the *how* of human creativity and what it means for the technical trajectory of LLMs, but this stimulus-free, unbounded, yet appropriate use of language is absent. More concerning, outside of their work, there is a possible conflation of “autonomous” AI systems with CALU, without realizing that this is what *human* autonomy entails.

Peter Voss and Mladan Jovanović (2023), for example, argue that some of the reasons why artificial general intelligence (AGI) has not yet been created is that current state-of-the-art systems lack the ability to autonomously learn (objects, actions, sequences, and appropriate contexts) and form new concepts and do so in real-time over the entire lifetime of the agent (Voss and Jovanović, 2023, 3-4). This echoes the reliance on *prompting* by end-users that hinders the level of creativity achieved by LLMs on Boden’s taxonomy, as argued by Francescella and Musolesi (2023, 5).<sup>15</sup>

It is worthwhile to ask, given that human creativity is the best baseline available for computational creativity, where CALU begins and ends in the construction of future AI systems that escape this reliance on prompting and begin to exhibit “autonomous” characteristics. Put another way: how much of what is described as “autonomy” in humans is actually CALU?

Questions for future research across disciplines become apparent:

---

<sup>15</sup> Voss and Jovanović (2023, 1) refer to this as a reliance on “external intelligence” that shifts the burden from the program to the programmer.

- (1) Exactly how much autonomy in future AI implicitly depends on a system being able to freely deploy its available conceptual resources in a manner that is not caused by its circumstances, done over an unlimited range, and yet appropriate to the circumstances?
- (2) Does a comprehensive definition of (actual or perceived) computational creativity implicitly depend on CALU?
- (3) Can CALU be scientifically explained?
- (4) If CALU permanently lacks scientific explanation, does this mean it cannot realistically be engineered into future AI systems?
- (5) Floridi and Chiriatti argue that the ability to "*prompt & collate*" will become critical for writers, representing a "different application of [human intelligence]" (Floridi and Chiriatti, 2020, 691). In contrast to inventions like the printing press, though, does the human reliance on LLMs effectively stifle CALU rather than allow it to unfold organically?
- (6) How much do stimulus-*free* linguistic productions—that are then brought into a broader human context in an unbounded fashion with the use of technologies—matter for human creativity?

These are open-ended questions, as some levels of creativity and other capabilities are clearly possible without CALU, and new AI models have had unexpected creative uses already. They are worth seriously considering as research proceeds rapidly.<sup>16</sup>

The value of generative linguistic theory, as illustrated above, in this respect is to establish the intellectual means by which human cognitive systems can be identified amidst myriad interaction effects—the *how* of human creativity in relation to the *what*. While it is conceivable that future AI systems—governed by deep learning algorithms or not—may find different ways of achieving CALU, this is an odd thing to simply hope for based on current trends. Careful attention to how humans behave, and the ways in which this has been studied by generative linguists, is useful for the future of computational creativity.

## 7 Conclusion

The creative aspect of language use, we have seen, is the stimulus-free, unbounded, and appropriate use of language by human beings. While these three properties may be found in the application of human cognitive capacities outside of language, it is most strikingly found within this domain. We have seen how CALU relates to the *how* of creativity, in this way playing a central role in the creative *process*. This central role supports the intellectual freedom of movement one finds in the human arts and sciences, presenting a baseline account of creativity from which computational creativity can be evaluated. From this baseline account, we evaluated state-of-the-art LLM-powered systems and found them to possess only one criterion—unboundedness—therefore judging them to not exhibit CALU.

Despite this, a new conversation has been inaugurated in part by the rise of generative LLMs demonstrating limited but surprising capabilities and emerging scholarly work (e.g., Franceschelli and Musolesi, 2023) on the intersection of human and computational creativity. Cognitive science and philosophy remain, in a number of ways, crucially relevant to

---

<sup>16</sup> See, e.g., Park et al. (2023) for research on a simulation of autonomous agents underpinned primarily, but not exclusively, by generative LLMs. I have personal doubts about the utility of this project, but it demonstrates enthusiasm for autonomous agents.

understanding both and, potentially, contributing to the development and practical application of systems capable of the latter. With a thorough explication and defense of CALU against this backdrop, it is high time for this ability to form part of a baseline account of human creativity and one that should inform our evaluations of and expectations for computational creativity going forward.

## References

- Alignment Research Center. (2023, March 17). *Update on ARC's recent eval efforts*. ARC Evals. Retrieved September 21, 2023, from <https://evals.alignment.org/blog/2023-03-18-update-on-recent-evals/>
- Allott, N., Lohndal, T., & Rey, G. (2021). Chomsky's "Galilean" explanatory style. In N. Allot, T. Lohndal, & G. Rey (Eds.), *A companion to Chomsky* (pp. 515-528). Wiley.
- Anthropic. (2023, March 14). *Introducing Claude*. Anthropic. Retrieved September 21, 2023, from <https://www.anthropic.com/index/introducing-claude>
- Asoulin, E. (2013). The creative aspect of language use and the implications for linguistic science. *Biolinguistics*, 7, 228-248. <https://doi.org/10.5964/bioling.8963>
- Baker, M. C. (2008). The creative aspect of language use and nonbiological nativism. In P. Carruthers, S. Laurence, and S. Stich (Eds.), *The Innate Mind, Volume 3: Foundations and the future* (pp. 233-253). Oxford University Press.
- Barone, J. (2023, May 25). *The 'Succession' soundtrack is fit for a concert hall*. The New York Times. Retrieved September 15, 2023, from <https://www.nytimes.com/2023/05/25/arts/television/succession-soundtrack-classical-music.html>.
- Boden, M. A. (2004). *The creative mind: myths and mechanisms*, Second Edition. Routledge.
- Boden, M. A. (2006). *Mind as machine: a history of cognitive science*, Volume 1 & 2. Oxford University Press.
- Boden, M. A. (2009). Computer models of creativity. *AI Magazine*, 30(3), 23-34. <https://doi.org/10.1609/aimag.v30i3.2254>
- Briskman, L. (1981). Creative product and creative process in science and art. In D. Dutton and M. Krausz (Eds.), *The concept of creativity in science and art* (pp. 129-155). Martinus Nijhoff Publishers.
- Chomsky, N. (1964). *Current issues in linguistic theory*. De Gruyter.

- Chomsky, N. (1967). Recent contributions to the theory of innate ideas. *Synthese*, 17(1), 2-11. <https://doi.org/10.1007/BF00485013>
- Chomsky, N. (1989). *Language and problems of knowledge: the Managua lectures*, Third Printing. The MIT Press.
- Chomsky, N. (2006). *Language and mind*, Third Edition. Cambridge University Press.
- Chomsky, N. (2007). Symposium on Margaret Boden, Mind as machine: a history of cognitive science, Oxford (2006), two volumes. *Artificial Intelligence*, 171(18), 1094-1103. <https://doi.org/10.1016/j.artint.2007.10.005>
- Chomsky, N. (2009a). *Cartesian Linguistics: A chapter in the history of rationalist thought*, Third Edition. Cambridge University Press.
- Chomsky, N. (2009b). The mysteries of nature: how deeply hidden? *The Journal of Philosophy*, 106(4), 167-200. <https://doi.org/10.7312/bric14474-002>
- Chomsky, N. (2009c). Turing on the “Imitation Game.” In R. Epstein, G. Roberts, & G. Beber (Eds.), *Parsing the Turing Test: Philosophical and methodological issues in the quest for the thinking computer* (pp. 103-106). Springer Netherlands.
- Cohen, I. B. (1985). *Revolution in science*. Belknap Press of Harvard University Press.
- Collins, J. (2023). Generative linguistics: ‘Galilean style.’ *Language Sciences*, 100, 1-14. <https://doi.org/10.1016/j.langsci.2023.101585>
- D’Agostino, F. (1984). Chomsky on creativity. *Synthese*, 58(1), 85-117. <https://doi.org/10.1007/BF00485363>
- Descartes, R. (1910/1637). *Discourse on the method of rightly conducting the reason, and seeking truth in the sciences*. The Open Court Publishing Company.
- Drach, M. (1981). The creative aspect of Chomsky’s use of the notion of creativity. *The Philosophical Review*, 90(1), 44-65. <https://doi.org/10.2307/2184372>.
- Floridi, L. & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(2), 681-694. <https://doi.org/10.1007/s11023-020-09548-1>
- Franceschelli, G. & Musolesi, M. (2023). On the Creativity of Large Language Models. *ArXiv*. 1-17. <https://doi.org/10.48550/arXiv.2304.00008>
- Gaut, B. (2003). Creativity and imagination. In B. Gaut & P. Livingston (Eds.), *The creation of art: new essays in philosophical aesthetics* (pp. 148-173). Cambridge University Press.
- Harman, G. Book Review: Cartesian linguistics: a chapter in the history of rationalist thought. *The Philosophical Review*, 77(2), 229-235. <https://doi.org/10.2307/2183323>
- Helmore, E. (2023, May 28). *Absurdity meets gravitas: how Nicholas Britell scored Succession*. The Guardian. Retrieved September 15, 2023, from <https://www.theguardian.com/tv-and-radio/2023/may/28/absurdity-meets-gravitas-how-nicholas-britell-scored-succession#:~:text=Britell%20said%20he%20was%20inspired,An%20Open%20Letter%20to%20NYC..>

- Joly, A. (1985). Cartesian or Condillacian linguistics? *Topoi*, 4(2), 145-149.  
<https://doi.org/10.1007/BF00135841>
- Kodner, J., Payne, S., & Heinz, J. (2023). Why linguistics will thrive in the 21<sup>st</sup> century: a reply to Piantadosi (2023). *ArXiv*. 1-28. <https://doi.org/10.48550/arXiv.2308.03228>
- Land, S. K. (1974). The Cartesian language test and professor Chomsky. *Linguistics*, 12(122), 11-24. <https://doi.org/10.1515/ling.1974.12.122.11>
- Leivada, E., Murphy, E., & Marcus, G. (2023). DALL·2 fails to reliably capture common syntactic processes. *Social Sciences & Humanities Open*, 8(1), 1-10.  
<https://doi.org/10.1016/j.ssaho.2023.100648>
- Lobina, D. J. (2017). *Recursion: A computational investigation into the representation and processing of language*. Oxford University Press.
- Maher, M. L. (2010). Evaluating creativity in humans, computers, and collectively intelligent systems. *Proceedings of the 1st DESIRE Network Conference on Creativity and Innovation in Design*. 22-28. <https://dl.acm.org/doi/10.5555/1854969.1854977>
- McGilvray, J. (2001). Chomsky on the creative aspect of language use and its implications for lexical semantic studies. In F. Busa & P. Bouillon (Eds.), *The language of word and meaning* (pp. 5-27). Cambridge University Press.
- McGilvray, J. (2005). Meaning and Creativity. In J. McGilvray (Ed.), *The Cambridge companion to Chomsky* (pp. 204-222). Cambridge University Press.
- McGilvray, J. (2009). *Introduction to Cartesian Linguistics: A chapter in the history of rationalist thought*. Cambridge University Press.
- McGilvray, J. (2014). *Chomsky: language, mind, and politics*. Wiley.
- McGilvray, J. (2017). Cognitive science: what should it be? In J. McGilvray (Ed.), *The Cambridge companion to Chomsky* (pp. 175-195). Cambridge University Press.
- Microsoft. (2023). *Introducing the new Bing: the AI-powered assistant for your search*. Microsoft Edge. Retrieved September 21, 2023, from <https://www.microsoft.com/en-us/edge/features/the-new-bing?form=MT00D8>
- Mitchell, M. (2023, June 12). *Did GPT-4 hire and then lie to a Task Rabbit worker to solve a CAPTCHA? AI: A guide for thinking humans*. Retrieved September 13, 2021, from <https://aiguide.substack.com/p/did-gpt-4-hire-and-then-lie-to-a>
- Mitchell, M. & Krakauer, D. C. (2023). The debate over understanding in AI's large language models. *PNAS*, 120(13), 1-5. <https://doi.org/10.1073/pnas.2215907120>
- OpenAI. (2022, November 30). *Introducing ChatGPT*. OpenAI. Retrieved September 21, 2023, <https://openai.com/blog/chatgpt>
- OpenAI. (2023). GPT-4 technical report. *ArXiv*. 1-100.  
<https://doi.org/10.48550/arXiv.2303.08774>

- Park, J. S., O'Brien, J. C., Cai, J. C., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: interactive simulacra of human behavior. *ArXiv*. 1-22.  
<https://doi.org/10.48550/arXiv.2304.03442>
- Piantadosi, S. (2023). Modern language models refute Chomsky's approach to language. *Lingbuzz*, 1-57. [lingbuzz/007180](https://lingbuzz/007180).
- Pichai, S. (2023, February 06). *An important next step on our AI journey*. Google. Retrieved September 21, 2023, from <https://blog.google/technology/ai/bard-google-ai-search-updates/>
- Rhodes, M. (1961). An analysis of creativity. *The Phi Delta Kappan*, 42(7), 305-310.  
<https://www.jstor.org/stable/20342603>
- Ritchie, G. (2007). Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17(1), 67-99. <https://doi.org/10.1007/s11023-007-9066-2>
- Root, M. (1975). Language, rules, and complex behavior. In K. Gunderson (Ed.), *Language, mind, and knowledge* (Minnesota Archive Editions, pp. 321-343). University of Minnesota Press.
- Sampson, G. (2016). Two Ideas of Creativity. In M. Hinton (Ed.), *Evidence, experiment and argument in linguistics and the philosophy of language* (Peter Lang Edition, pp. 15-26). Peter Lang.
- Shanahan, M. (2023). Talking About Large Language Models. *ArXiv*. 1-13.  
<https://doi.org/10.48550/arXiv.2212.03551>
- Treisman, R. (2023, May 25). *How composer Nicholas Britell created the sound of 'Succession.'* NPR. Retrieved September 15, 2023, from <https://www.npr.org/2023/05/25/1178008172/nicholas-britell-succession>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460.  
<https://doi.org/10.1093/mind/LIX.236.433>
- Voss, P. & Jovanović, M. (2023). Why We Don't Have AGI Yet. *ArXiv*. 1-8.  
<https://doi.org/10.48550/arXiv.2308.03598>