Texts Without Authors:

Ascribing Literary Meaning in the Case of AI

Abstract: With the increasing popularity of Large Language Models (LLMs), there has been an increase in the number of AI generated literary works. In the absence of clear authors, and assuming such works have meaning, there lies a puzzle in determining who or what fixes the meaning of such texts. I give an overview of six leading theories for ascribing meaning to literary works. These are Extreme Actual Intentionalism, Modest Actual Intentionalism (1 & 2), Conventionalism, Actual Author Hypothetical Intentionalism, and Postulated Author Hypothetical Intentionalism. I argue that while only Conventionalism and Postulated Author Hypothetical Intentionalism show any promise of adjudicating how we ought to ascribe meaning in the case of AI generated texts, Postulated Author Hypothetical Intentionalism is the stronger of the two views.

1. The Case of AI

While the question of who or what fixes the meaning of literary texts is not new, developments in Large Language Models (LLMs) both complicate and illuminate the question. If we understand some AI-generated works as *texts*, and such texts have *meaning*, then there is a puzzle in determining who or what fixes said meaning. Views that defer to authorial intent are of little help, given the absence of a clear *author* in these cases. I argue that such texts can and do acquire meaning irrespective of the presence of an author.

This paper proceeds as follows. First, I provide a brief overview of how LLMs—such as GPT-4—work on a technical level. I then draw a tripartite distinction between *Illusory Texts, Artificial Texts, and Traditional Texts*. I argue that some Artificial Texts have fixed meanings. Then, I give an overview of six popular views on how the meaning of texts are fixed. I demonstrate that only two of the views, *conventionalism* and *postulated author hypothetical intentionalism*, show any promise of solving the puzzle. I argue against conventionalism and in

favour of postulated author hypothetical intentionalism. I conclude by considering some implications for our understandings of texts generated by AI.

2. Large Language Models

Contemporary LLMs are a kind of *generative AI* that produce text outputs based on user inputs in the form of *prompts*. Contemporary LLMs are *Deep Learning Models* which are *neural networks* with many layers. A neural network is an interconnected structure of nodes—called *neural units*—that process input and generate output. LLMs are trained on large *datasets* of text scraped from the internet, taken from books, etc. Words are then encoded into *word embeddings*, vector representations of words in a multidimensional space, with related words clustering together (Stinson 2024). For example, if one were to input the prompt "What is the capital of Canada?" the model would recognize "capital" and "Canada" as existing syntactically in relation to one another and would likely generate "Ottawa" as output.

Many contemporary LLMs use a model architecture called a *Transformer* (Radford et al. 2018; 2019; Brown et al. 2020; Ouyang et al. 2022). First proposed by Vaswani et al (2017), the Transformer architecture is a neural network wherein each unit uses a mechanism called *self-attention*, which is capable of "relating different positions of a single sequence in order to compute a representation of the sequence" (2). Rather than analyse the elements of a sentence in a linear fashion, LLMs built using Transformer architecture are able to analyse the different elements of a sentence nonlinearly. This provides stronger results in text generation, as more relationships between word embeddings can be considered concurrently, rather than processed one at a time.

For example, if one were to ask an application like ChatGPT "What does a cat like to eat?" a simple response would be "food" as the words "eat," and "food" are often found in close proximity to each other in many English language sentences. However, the additional context of "cat" and "likes" might provide a more sophisticated response, such as "fish," or "tuna." LLMs are thus able to generate sentences based on both the likelihood of certain words being in close proximity to one another in the training data and the likelihood of certain word embeddings being clustered together. This is why prompts such as "Write me a story about a knight slaying a dragon in the style of Edgar Allen Poe" often yield satisfying results. The model is able to recognize both Poe's vocabulary and syntax, as well as the tropes surrounding fantasy stories about knights slaying dragons.

Given that LLMs work primarily by pattern recognition, it is not as though they have *intentional attitudes* toward the text that they generate.[1] As such, they cannot be properly considered *authors* in the sense of being the intentional meaning-makers of texts. Likewise, the prompt-giver could not be considered an author proper, as they do not generate any of the text, nor do they know in advance what the content of the text output will be.

There are two possible moves one could make at this juncture. The first move would be to deny that such outputs can be considered texts proper, possessing neither an author nor literary meaning. The second move would be to accept the outputs as texts, but to deny that they have a fixed literary meaning. The former move would come from thinkers following Gracia (1995; 1996), whereas the latter might come from a follower of Barthes ([1967] 1977). In the next two sections, I provide an argument against each in turn, ultimately endorsing the view that AI-generated works can be texts, and that such texts have fixed meanings.

3. Three Kinds of Texts

One way to solve the problem of ascribing meaning to AI-generated works is to deny that such works can be literary texts.[2] There are several reasons why we should not take this to be the case.

We might consider works like David Jhave Johnston's *ReRites*, a collection of AI-generated poetry. *ReRites* is a twelve volume poetry set compiled between 2017 and 2018 consisting of AI-generated output arranged and edited by Johnston and later released in an edited format alongside critical essays (2019). Elsewhere, Allison Parish's *Articulations* (2018) takes a similar approach by training an LLM on a dataset of public domain poetry and publishing a curated set of results. Works such as these—and *electronic literature* more broadly—have been taken up in academic circles among literature scholars as both subjects of criticism (MacBeth 2023) and texts that challenge traditional understandings of literature (Hayles 2018).

ReRites and *Articulations* were presented as literary works and taken up as such. This is in line with institutional theories of art such as those by Danto (1964) and Dickie (1969; 1974) that hold for a given work to be a work of art, it must be presented to an *artworld*. Levinson (1996) holds a similar view about literary works. According to institutional theories of art, a work like *ReRites* would indeed be a work of literature. We might also consider the move in contemporary literature to consider any work a potential text (Goldsmith 2011). According to conceptual poets, even a phonebook can be a literary text if it is presented as such.

This line of argumentation would likely not persuade those in Gracia's (1995; 1996) camp. For Gracia, a text must contain an intentional element. In order for a work to constitute a text proper, it must be the product of an intentional agent. Shakespeare's *King Lear* is a text, a cluster of rocks randomly arranged by the wind to spell out a verse of Shakespeare is not. However, something strange occurs when an artist takes a photo of the rocks, labels them

"Shakespeare Rocks" and presents the work to an artworld.[3] If the work is taken up by audiences and critics, it would have a claim to text status.[4]

There seems then, to be a distinction between *King Lear*, the cluster of rocks, and "Shakespeare Rocks." We might draw the lines this way: *King Lear* is a *Traditional Text*. It is the product of an intentional agent—an author—who decided to create the work. The cluster of rocks is an *Illusory Text*. In other words, while it might have the appearance of a text, it is no text at all. "Shakespeare Rocks" appears to occupy a strange middle ground. While it was not created by an intentional agent, its status as a literary work depends on an intentional agent. The same goes for AI-generated works. They cannot be said to have been created by an intentional agent, but their creation is dependant on the agency of a curator. In this sense we might think of works like "Shakespeare Rocks" and *ReRites* as *Artificial Texts*—texts without authors.

One could argue that in the case of "Shakespeare Rocks" the photographer is the author of the work, given that we consider photographers to be the authors of the pictures they take in general. Here I make the distinction between "Shakespeare Rocks" qua photograph and "Shakespeare Rocks" qua text. While it is true that the photographer is the author of "Shakespeare Rocks" qua photograph, they are not the author of "Shakespeare Rocks" qua text. Here it is useful to consider Christy Mag Uidhir's (2011) notion of x being the author of a work W as an F. Consider the example of films. The writer of a film is the author of the film as a script, whereas the director of a film is the author of the film as a film. Similarly, the photographer is the author of "Shakespeare Rocks" as a photograph, but not as a text.

Furthermore, the photographer did not imbue the text with meaning, they noticed something interesting about the rock formation (that it spells out a verse of Shakespeare) and presented it to an audience for consideration. The photographer acts as a curator in this scenario

rather than an author. Cases like *Articulations* are similar, insofar as the text is produced by AI and then taken up by a curator and presented to an audience. "Shakespeare Rocks," *Articulations*, and *ReRites* possess interesting textual meaning irrespective of a curator. However, it is the curator that transforms the text from an Illusory Text to an Artificial Text.[5] The curators in these examples are inviting the audience to consider these works as texts, not as photographs or other artforms. Left alone, "Shakespeare Rocks" would remain an Illusory Text. However, it becomes an Artificial Text when a curator presents it as such. The role of the curator differs from the role of the author. In this way we can consider Artificial Texts to be texts without authors.

In summary, we can think of three distinct kinds of text. First, there are Illusory Texts, which take on the appearance of texts but are not texts themselves. Then, there are Traditional Texts, which are the products of intentional agents. Finally, there are Artificial Texts, which are not the products of intentional agents, but whose status as a text depends on the agency of a curator. AI-generated texts are thus a different kind of text than Traditional Texts, yet we can think of them as texts, nonetheless.

However, despite their status as texts, it could still be the case that Artificial Texts—such as AI-generated works or found poetry—do not have fixed meanings. I address this position in the next section.

4. Meaning and Artificial Texts

In this section I pursue the question as to whether some Artificial Texts have fixed meanings. Before addressing this, however, I will first define *literary meaning* and offer three theories for understanding literary meaning.

I take literary meaning to be distinct from semantic meaning insofar as a work of literature can come to mean more than the sum of its atomic parts. For example, a literal reading of Orwell's *Animal Farm* reveals a story about animals on a farm. However, a literary reading of the same text reveals an allegorical work about political struggle. The semantic meaning is the former, whereas the literary meaning is the latter. According to Olsen, "…it has become a critical commonplace that the literary work is a verbal expression. a verbal construct or an utterance, and that as such its peculiar nature is defined through the special *way in which it means*" (1982, 13). On this account, literary meaning is not reducible to semantic meaning, though semantic theories inform literary ones.

Olsen (1982) identifies three kinds of theory of literary meaning: there are *autonomy theories*, *semiotic theories*, and *intentionalist theories*. Autonomy theories (Frye 1957; Beardsley 1958; Richards 1965) hold that the meaning of a text is immanent to the text. Semiotic theories (Propp 1968; Culler 1981) hold that meaning is determined through the interpretation of the relationship between the signs and systems that govern a given text.[6] Intentionalist theories (Hirsch 1967; Juhl 1980) hold that the meaning of a text is imbued by the intention of the author of the text.[7] For both autonomous and intentionalist theories, a given text has a fixed literary meaning. Semiotic theories differ as to whether they consider texts to have fixed meanings. Of those that do not hold that texts have fixed meanings, Barthes ([1967] 1977) is most notable.

For Barthes, the author of a text does not hold a privileged position with respect to the meaning-making of a text. Rather, the meaning of a text is determined collectively by the audience—of whom the author is a part. This can be read as conservatively as there being one meaning generated collectively by the audience, or as radically as there exists a different

meaning for every reader. It might be more accurate in the second case to say that the text does not possess a meaning at all, rather there are only audience interpretations.

There is a major problem with this view however, as we can be mistaken about what a text means. In light of this, some interpretations of a text are more correct than others. Consider *Animal Farm*. It would be incorrect to interpret the text as being an allegory for the 2016 US election, as the writing and publication of *Animal Farm* greatly predates the election.[8] One could argue that over time texts acquire new meanings. However, this ignores the relevance of the context in which a text was produced and generates bizarre ahistorical readings.

It is more reasonable to interpret *Godzilla* as a metaphor for nuclear weapons than it is to interpret the H.G. Wells' *War of the Worlds* as one. This is because nuclear weapons did not exist at the time Wells was active and publishing. While we can *interpret* texts in a variety of ways, that does entail that such interpretations correspond to *meanings*. If the goal of interpretation is to discern the meaning of a text, then not all interpretations of a text are good interpretations as they do not fulfill this function.[9] It does not follow from this that all Artificial Texts have fixed meanings. It is very likely that many do not.[10]

According to literary critic Gérard Genette (1992) *transtextuality* is everything that brings a text into relation, either explicitly or implicitly with other texts. Under the umbrella of transtextuality is *metatextuality*, that which connects a commentary of a text to the text itself. Artificial Texts are often conceptual in nature. As such, much of the meaning one is to find in such texts exists through their meta level commentary. This can be commentary on other texts, or on the text in question. In what follows I give two examples of Artificial Texts that have fixed meanings at the metatextual level.

First, consider *Via: 48 Dante Variations* by Caroline Bergvall. As a work of found poetry, *Via* is a compilation of 48 iterations of the first canto of Dante's *Inferno*. It begins with the original Italian and follows with 47 unique English translations. The work explores the process of translation and what is lost and gained through the act. A direct reading of any individual part of the text does not reveal its meaning, however, by reading the text as a whole—with the various translations juxtaposed—one can come to understand the work as being about translation.

Second, consider *Death of an Author* by Stephen Marche. *Death of an Author* is a detective novella written almost entirely with the use of models like ChatGPT. The story follows literary critic Gus Dupin as he investigates the murder of celebrated novelist Peggy Firmin. Dupin becomes embroiled in a twisted plot involving the LLM company Marlow AI as he himself becomes a murder suspect. While on the surface a work of detective fiction, the novel explores the relationship between humans and AI, especially as they pertain to the realm of literature. In that sense the novella is a mirror of its own creation and acts as a metaphor for its own process.

Both *Via* and *Death of an Author* are Artificial Texts with what appear to be fixed meanings, with the former being about translation and the latter about human-AI relations. Granted that both texts have fixed meanings, then the puzzle of who or what fixes the meaning of said texts becomes apparent. In the next section I outline six competing views on attributing literary meaning, and jettison four as viable contenders.

5. Six Views of Ascribing Literary Meaning

According to Irvin (2006), there are six leading theories of how a literary text's meaning is fixed.[11] These are extreme actual intentionalism (EAI), two versions of modest actual intentionalism (MAI1 and MAI2), conventionalism, actual author hypothetical intentionalism (AAHI) and postulated author hypothetical intentionalism (PAHI).[12] I will go over each in brief detail before dismissing EAI, MAI1, MAI2, and AAHI as viable solution bearers.

Extreme actual intentionalism (Irwin 1999; 2015; Stock 2017; García-Carpintero 2019; 2023) is the view that it is the author's intent that fixes the meaning of a text, irrespective of linguistic conventions. This means that if an author intended for the sentence "the apple is red" to mean "the banana is yellow," then we ought to take the second to be the meaning of the first in that context. Modest actual intentionalism (Carroll 2000) deals with this strange feature of EAI by taking the view that when an authors intent coincides with linguistic convention, it is the author's intent that fixes the meaning. Proponents of MAI1 hold that when the author's intent diverges from linguistic convention, the affected portion of the text is meaningless. Proponents of MAI2 hold that when this occurs, either meaning is fixed by linguistic convention, or that portion of the text remains ambiguous (Stecker 2003). MAI2 can collapse into conventionalism if a given text is especially ambiguous.

Conventionalism holds that it is linguistic convention that fixes the meaning of a text (Beardsley 1958), although some have argued that context ought to be considered as well (Davies 2005). Conventionalists disagree over what to do in cases where the text is ambiguous, but do not defer to an author in such cases. Hypothetical intentionalism comes in two forms. Actual author hypothetical intentionalism holds that the meaning of a given work is fixed by the best hypotheses an informed audience would form about the actual author's intentions (Levinson 1996). Postulated author hypothetical intentionalism differs insofar as it is not the actual author

of a work, but an ideal author that possesses additional features such as full mastery of language (Lin 2018; 2023).[13] AAHI often collapses into a version of actual intentionalism, while PAHI runs the risk of collapsing into conventionalism. However, as I will later argue, there are cases where the distinction between conventionalism and PAHI becomes meaningful such that we might prefer PAHI.

In the case of AI-generated works, it does not make sense to defer to an author, as Artificial Texts are *texts without authors*. This fact rules out EAI, MAI1, MAI2, and AAHI as contenders for a theory of ascribing meaning that can solve the puzzle of AI-generated texts, given that they all require the presence of an author for meaning to be fixed in at least some cases. For EAI and AAHI, there is simply no author to whom one could look to for meaning. In MAI1, the AI-generated text in question would have to be conceived as meaningless—I have already given cases where we should not take this to be true. In MAI2, the meaning of the text would be fixed by linguistic convention, which is to say that in cases of AI-generated works, MAI2 *always* collapses into conventionalism.

There is one additional move that defenders of EAI could make in an attempt to salvage the explanatory power of EAI in cases wherein the text in question is AI-generated. Irwin (2015) draws a distinction between *meaning* and *significance*. The former is the interpretation of a text intended by its author. The latter is any interpretation of a text that differs from that intended by its author. Queer readings of literature that lacks explicit reference to queerness would count as examples of significance. On this account it follows that in the absence of an author, AI-generated text has no fixed meaning but many legible readings with significance.

I have already argued that some Artificial Texts have fixed meanings, however one need not turn to AI-generated text for examples of texts without fixed meanings. Consider the song “I

Am the Walrus" by *The Beatles*. The song was written without the explicit intent of having a meaning, and yet music critics and audiences alike have drawn their own saliant interpretations of the song's nonsense lyrics. We might also consider the art movement dadaism as a way of making texts without fixed meaning that still possess significance. However, there is a huge leap from the nonsense lyrics of "I Am the Walrus" or dadaist sound poetry to AI-generated text. While the former two seem to be examples of intentional meaninglessness, the latter appears to be the opposite: unintentional meaningfulness. Both "I Am the Walrus" and dadaist poems feature odd word choice and deeply unusual syntax and semantics.

Popular LLMs like ChatGPT seem to be doing something different insofar as the syntax and semantics of AI-generated sentences utilize linguistic conventions to produce meaningful content. There is a stark difference between the nonsense phrase "goo goo g'joob" and a work such as *Death of an Author*. The former lacks obvious meaning, whereas the latter contains obvious meaning. Reducing the meaning of a text like *Death of an Author* to significance seems to be a case of linguistic confusion. If there is only one significance of a text, it would make more sense to consider that the meaning of said text, rather than to maintain that the text in question has no meaning. In light of this, EAI remains inappropriate as a candidate for explaining meaning in cases of AI-generated texts.

While I have ruled out four of the six competing theories of ascribing meaning to a text, conventionalism and PAHI remain on the table. In the next section I argue that we ought to reject conventionalism—despite its seeming potential to resolve the puzzle—and instead embrace PAHI

6. Fixing the Meaning of Artificial Texts

While at first glance conventionalism seems as though it can adequately solve the puzzle of attributing meaning to AI-generated texts—by holding that it is *only* linguistic convention that fixes meaning, thus circumnavigating the problem of the absent author—it runs into problems insofar as it is unable to account for meaning found at a metatextual level.

Consider the two Artificial Texts discussed earlier, *Via* and *Death off an Author*. There is nothing about the individual words or passages of *Via* that would lead to the interpretation that the text is about translation. Rather, the text must be taken holistically. It is only by understanding *Via* as being constructed by juxtaposing different translations of Dante can we understand it as a work about translation. *Death of an Author*, on the other hand, *can* be interpreted through the individual parts of the text. However, the larger metatextual point of the work itself being imbricated in its meaning—it is an exemplar of the themes of the narrative—is lost.

Thus far I have identified two main problems in ascribing meaning in the case of AI-generated texts. The first comes with most forms of intentionalism that rely on the presence of an author for a text to acquire meaning. In the case of AI, there is no author to which we can defer. The second problem lies with conventionalism's inability to account for meaning on a metatextual level. I will show how PAHI is able to effectively deal with both of these problems.

First, the problem of authorship. According to PAHI, it is not the author who fixes the meaning of a text. Rather, it is the audience's best hypotheses, considering relevant evidence, of what meaning an ideal author—who possesses such features as mastery of the English language—would imbue within the text that fixes meaning. As it is the audience that fixes

meaning with their informed hypotheses given the supposition of an ideal author, the problem of an absent author is circumnavigated entirely. We posit an author—call them Stephen Marche — that possesses mastery of style and language in addition to having reasonable intentions surrounding the meaning of their work. In this way, the postulated author Stephen Marche serves a role in fixing the meaning of AI-generated text in the absence of an actual author Stephen Marche.

Second, given the postulation of an ideal author, PAHI can account for metatextual meaning in ways conventionalism cannot. The postulated author of PAHI possesses mastery of language, including the use of metatextual writing. Rather than understand the meaning of a text to be merely the sum of individual components, PAHI can consider texts holistically to unlock their meanings at the metatextual level.

While both conventionalism and PAHI can deal with the first problem—the absence of an author, only PAHI can deal with the second—how texts can come to have meaning at a metatextual level. Given this, we ought to accept PAHI as being able to account for the fixing of meaning in AI-generated texts.

7. Consequences

In this article I have argued three things. First, that AI-generated texts ought to be considered Artificial Texts—texts without authors. Second, that Artificial Texts sometimes have a fixed literary meaning. Third, the view that best explains how the literary meaning of Artificial Texts is fixed is Postulated Author Hypothetical Intentionalism.

There are two main consequences of my view. The first is that we now have the language of Artificial Texts to understand AI-generated works as texts proper. As Artificial Texts are

distinct from Traditional Texts, we know that the same rules and implications need not apply to both kinds of text. This distinction also helps explain how works such as found texts can be considered texts proper.

The second consequence is that once we determine an Artificial Text has a fixed meaning, we have an interpretive framework—PAHI—for discerning said meaning. The difficulty now lies in constructing a rubric or heuristic for determining which Artificial Texts have fixed literary meanings, and which Artificial Texts may be (re)interpreted by the audience ad nauseum.

References

Barthes, Roland. (1967) 1977. “The Death of the Author.” In *Image, Music, Text*, translated by Stephen Heath, 142–48. London: Fontana.

Beardsley, Monroe C. 1958. *Aesthetics: Problems in the Philosophy of Criticism*. New York: Harcourt, Brace.

Block, Ned. 1981. “Psychologism and Behaviorism.” *The Philosophical Review* 90 (1): 5–43. https://doi.org/10.2307/2184371.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. “Language Models Are Few-Shot Learners.” arXiv. https://doi.org/10.48550/arXiv.2005.14165.

Carroll, Noël. 2000. “Interpretation and Intention: The Debate Between Hypothetical and Actual Intentionalism.” *Metaphilosophy* 31 (1/2): 75–95.

Culler, Jonathan D. 1981. *The Pursuit of Signs: Semiotics, Literature, Deconstruction*. Ithaca, N.Y: Cornell University Press.

Currie, Gregory. 1990. *The Nature of Fiction*. Cambridge [England] ; Cambridge University Press.

Danto, Arthur. 1964. “The Artworld.” *The Journal of Philosophy* 61 (19): 571–84.

Davies, Stephen. 2005. “Beardsley and the Autonomy of the Work of Art.” *The Journal of Aesthetics and Art Criticism* 63 (2): 179–83.

Dickie, George. 1969. “Defining Art.” *American Philosophical Quarterly* 6 (3): 253–56.

———. 1974. *Art and the Aesthetic : An Institutional Analysis*. Ithaca, NY: Cornell University Press.

Frye, Northrop. 1957. *Anatomy of Criticism: Four Essays*. Princeton, NJ: Princeton University Press.

García-Carpintero, Manuel. 2019. “On the Nature of Fiction-Making: Austin or Grice?” *The British Journal of Aesthetics* 59 (2): 203–10. https://doi.org/10.1093/aesthj/ayy054.

———. 2023. “The Semantics of Fiction.” *Mind & Language* 38 (2): 604–18. https://doi.org/10.1111/mila.12412.

Genette, Gérard. 1992. *The Architext: An Introduction*. Quantum Books. Berkeley: University of California Press.

Goldsmith, Kenneth. 2011. *Uncreative Writing: Managing Language in the Digital Age*. New York City: Columbia University Press.

Gracia, Jorge. 1995. *A Theory of Textuality: The Logic and Epistemology*. Albany, NY: SUNY Press.

———. 1996. *Texts: Ontological Status, Identity, Author, Audience*. Albany, NY: State University of New York Press.

Hayles, Katherine. 2018. “Literary Texts as Cognitive Assemblages: The Case of Electronic Literature.” *Electronic Book Review*, August. https://electronicbookreview.com/essay/literary-texts-as-cognitive-assemblages-the-case-of-electronic-literature/.

Hirsch, E. D. 1967. *Validity in Interpretation*. New Haven: Yale University Press.

Irvin, Sherri. 2006. “Authors, Intentions and Literary Meaning.” *Philosophy Compass* 1 (2): 114–28.

Irwin, William. 1999. *Intentionalist Interpretation: A Philosophical Explanation and Defense*. Westport, Conn.: Greenwood Press.

———. 2015. “Authorial Declaration and Extreme Actual Intentionalism: Is Dumbledore Gay?” *Journal of Aesthetics and Art Criticism* 73 (2): 141–47.

Johnston, David Jhave. 2019. *ReRites: Raw Output/Responses*. Montreal: Anteism Books.

Juhl, P. D. 1980. *Interpretation: An Essay in the Philosophy of Literary Criticism*. Princeton, N.J: Princeton University Press.

Levinson, Jerrold. 1996. “Intention and Interpretation in Literature.” In *The Pleasures of Aesthetics: Philosophical Essays*, 175–213. Ithaca, NY: Cornell University Press.

Lin, Szu-Yen. 2018. “Interpretation and the Implied Author: A Descriptive Project.” *The Southern Journal of Philosophy* 56 (1): 83–100. https://doi.org/10.1111/sjp.12269.

———. 2023. “Defending the Hypothetical Author.” *The British Journal of Aesthetics* 63 (4): 579–99. https://doi.org/10.1093/aesthj/ayac062.

MacBeth, George. 2023. “‘A Shifting Surface World’: The Techno-Graphomania of David Jhave Johnston’s ReRites</Em>.” *Electronic Book Review*, April. https://electronicbookreview.com/essay/a-shifting-surface-world-the-techno-graphomania-of-david-jhave-johnstons-rerites/.

Mag Uidhir, Christy. 2011. “Minimal Authorship (of Sorts).” *Philosophical Studies* 154 (3): 373–87. https://doi.org/10.1007/s11098-010-9525-0.

Nehamas, Alexander. 1981. “The Postulated Author: Critical Monism as a Regulative Ideal.” *Critical Inquiry* 8 (1): 133–49. https://doi.org/10.1086/448144.

Olsen, Stein Haugom. 1982. “The ‘Meaning’ of a Literary Work.” *New Literary History* 14 (1): 13–32. https://doi.org/10.2307/468955.

Ouyang, Long, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, et al. 2022. “Training Language Models to Follow Instructions with Human Feedback.” arXiv. https://doi.org/10.48550/arXiv.2203.02155.

Parrish, Allison. 2018. *Articulations*. Denver, CO: Counterpath.

Propp, Vladimir. 1968. *Morphology of the Folktale.* Second Edition. Bibliographical and Special Series of the American Folklore Society ; v. 9. Austin: University of Texas Press.

Radford, Alec, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. “Improving Language Understanding by Generative Pre-Training.” In . https://www.semanticscholar.org/paper/Improving-Language-Understanding-by-Generative-Radford-Narasimhan/cd18800a0fe0b668a1cc19f2ec95b5003d0a5035.

Radford, Alec, Jeff Wu, R. Child, D. Luan, Dario Amodei, and I. Sutskever. 2019. “Language Models Are Unsupervised Multitask Learners.” In . https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe.

Richards, Ivor Armstrong. 1965. *The Philosophy of Rhetoric*. The Mary Flexner Lectures, 1936. New York: Oxford University Press.

Searle, John. 1980. “Minds, Brains, and Programs.” *Behavioral and Brain Sciences* 3 (3): 417–24. https://doi.org/10.1017/S0140525X00005756.

Stecker, Robert. 2003. *Interpretation and Construction: Art, Speech, and the Law*. Malden, Mass.: Blackwell Publishing.

Stinson, Catherine. 2024. “What Are Large Language Models Made Of?” In *Leading the Way: Envisioning the Future of Higher Education*, edited by M. E. Norris and S. E. Smith.

Stock, Kathleen. 2017. *Only Imagine: Fiction, Interpretation and Imagination*. Oxford University Press. https://doi.org/10.1093/oso/9780198798347.001.0001.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. "Attention Is All You Need." arXiv. https://doi.org/10.48550/arXiv.1706.03762.

Walton, Kendall. 1990. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, Mass: Harvard University Press.

———. 2008. *Marvelous Images: On Values and the Arts*. Oxford: Oxford University Press.

Endnotes

[1] This claim about AI in general has been defended by others, perhaps most notably by John Searle (1980) and Ned Block (1981). I will defer to previous scholarship and not provide my own defense of the claim that AI lacks intentionality.

[2] Followers of Gracia (1995; 1996) would likely take this approach.

[3] Here I am playing with Walton's (1990) famous "Cracks in a Rock" thought experiment.

[4] Likewise with other found texts discussed by scholars like Goldsmith (2011).

[5] It might be the case that all Artificial Texts begin their existence as Illusory Texts.

[6] Texts can either be closed, such that no outside understandings of signs are permissible, or open, wherein the signs and systems that make up a text are in conversation with other discourses.

[7] I discuss these in greater detail in the next section.

[8] It remains an open question as to when and how a given text acquires its meaning. I discuss this issue in the next section.

[9] It is also possible, in principle, for a text to have multiple fixed meanings.

[10] This is not unique to Artificial Texts. There are Traditional Texts that appear to lack fixed meaning, such as James Joyce's Finnegans.Wake, a "novel" consisting only of surreal stream of consciousness word play and devoid of any obvious plot or characters.

[11] All of the theories presented rely on an intentionalist account of meaning, except for conventionalism, which relies on an autonomous account of meaning.

[12] I follow Irvin (2006) and Nehamas (1981) in labelling this author the postulated.author. Other names have been suggested, such as the apparent.artist (Walton 2008), the fictional.author (Currie 1990), and the hypothetical.author (Lin 2023). I take all of these terms to be more or less equivalent for my purposes.

[13] One might wonder about the difference between PAHI and death of the author (Barthes [1967] 1977). While in both cases it is the audience that fixes the meaning of a text, the postulated author of PAHI serves as a limiting function on the number and kinds of possible interpretations. If one agrees with Barthes about the death of the author, then one could plausibly argue that it is merely the audience that fixes the meaning of a text. As I hold that some interpretations of texts are more correct than others, I do not endorse this view.