NOTE · 22 JULY 2026

What is an Author?

4,340 WORDS · 19 MIN · FILED UNDER LITERARY THEORY

In the essay "Death of The Author" (https://writing.upenn.edu/~taransky/Barthes.pdf), Barthes argues for the disappearance of an origin for work—the removal of an individual which is the "Author"—and instead posits the idea of a structural evaluation of a work. However, his analysis can be found to be unsatisfactory in his disregard of an "Author" completely, and indeed, much of the work Foucault does in "What is An Author?" is to also consider the "Author" as a function in discourse and consider its properties while continuing to disregard the creative nature that traditional criticism [1] attributed to a work. In this manner, as he says, he wishes to examine this phenomenon of the disappearance of the "Author", through the idea of the "Author-function" and liberate himself once and for all, from this conceptual framework of the transcendental and critical tradition of the nineteenth century.

Since many of the arguments examine the author in the light of "discourses", it serves to know what he means specifically. From the Stanford Encyclopedia of Philosophy:

> Foucault explicitly presents The Order of Things as an "archaeological" approach to the history of thought. Three years later, in 1969, he published The Archaeology of Knowledge, a methodological treatise that explicitly formulates what he took to be the archaeological method that he used not only in The Order of Things but also (at least implicitly) in History of Madness and The Birth of the Clinic. The key idea of the archaeological method is that systems of thought and knowledge (epistemes or discursive formations, in Foucault's terminology) are governed by rules, beyond those of grammar and logic, that operate beneath the consciousness of individual subjects and define a system of conceptual possibilities that determines the boundaries of thought in a given domain and period. So, for example, History of Madness should, Foucault maintained, be read as an intellectual excavation of the radically different discursive formations that governed talk and thought about madness from the seventeenth through the nineteenth centuries.
>
> Archaeology was an essential method for Foucault because it supported a historiography that did not rest on the primacy of the consciousness of individual subjects; it allowed the historian of thought to operate at an unconscious level that displaced the primacy of the subject found in both phenomenology and in traditional historiography. However, archaeology's critical force was restricted to the comparison of the discursive formations of different periods. Such compar-

> isons could suggest the contingency of a given way of thinking by showing that the people living in previous ages had thought very differently (and, apparently, just as effectively). But mere archaeological analysis could say nothing about the causes of the transition from one way of thinking to another and so had to ignore perhaps the most forceful case for the contingency of entrenched contemporary positions. Genealogy, the new method first deployed in Discipline and Punish, was intended to remedy this deficiency.

In essence, he thinks about "discourses"—systems of thoughts and knowledge—as governed by subconscious rules pertaining to a (historical) period or the domain this knowledge is situated in.

PART I

Problems with the disappearance of the Author

First, he reminds us of the attributes concerning the (at that time period) understanding of the relationship between the author and their work. The work of Barthes and others introduce the "ethical" [2] principle of contemporary writing, wherein one recognises an indifference in *who* the one writing the text is. Writing is understood through its *interior* structure, and not by an external, originary source in the form of an "Author" who has inscribed some meaning to the text. The argument for this is linguistic in nature, writing is understood as an interplay of signs, where meaning is attained not from an external deployment of the sign (the physical form of the signifier) but rather, through an internal system of comparison where meaning is attained from the *difference* in signs.

He also points out how this has lead to a reversal of the nature of the relationship between writing and death. In classical tradition, a work is understood to have allowed an Author to escape their death in a sense, their psychological or creative nature (so their characteristics as an individual) would survive beyond their death through their works. However, in contemporary writing a reversal has happened—the works of, for example Proust and Kafka [3], have served to "kill" their author. The style of their work is such that it uses detached language, ensuring that the work removes all characteristics of the individual by whom it was written, letting the writing subject disappear into the language itself.

However, as we shall see, there is still much work to be done for understanding just what an author is. Foucault poses in particular, two problems with the (then) work pertaining to this phenomenon. One is the nature of "works". This is an empirical query, as he says, "If an individual is not an author, what are we to make of those things he has written or said, left among his papers or communicated to others? Is this not properly a work? What, for instance, were Sade's papers before he was consecrated as an author? [4] "

The question is not much easier if we are dealing with an "Author", is everything written down by him a work? (A similar line of thought, without necessarily appealing to an "Author" is in the previous footnote). What about works with no "individual" that substitutes as an author in the classical tradition, say *The Arabian Nights*? There are endless questions of this manner, the removal of an author does not bypass the problematic nature of this. Although Foucault does not talk about this, one may also point out that the analysis of Barthes, for instance, assumes the nature of the work as a unity of sorts, where the very nature of calling a collection of writing a "work" is supposed to smooth over its contradictions, and accord a singular, unified structure to it, much in the same way traditionally calling upon an author was supposed to smooth over all the contradictions in a work by pointing to a single unified conscious behind the text.

Another thesis that detains us from fully taking advantage of this new phenomenon is the notion of "écriture" or "arche-writing". Before I go into Foucault's critique of how it has served to only arrest the possibility of discarding traditional criticism, let us momentarily consider what "arche-writing" is supposed to be.

Here is the Internet Encyclopedia Of Philosophy entry on Derrida and arche-writing (écriture):

> In Of Grammatology and elsewhere, Derrida argues that signification, broadly conceived, always refers to other signs, and that one can never reach a sign that refers only to itself. He suggests that "writing is not a sign of a sign, except if one says it of all signs, which would be more profoundly true" (OG 43), and this process of infinite referral, of never arriving at meaning itself, is the notion of 'writing' that he wants to emphasise. This is not writing narrowly conceived, as in a literal inscription upon a page, but what he terms 'arche-writing'. Arche-writing refers to a more generalised notion of writing that insists that the breach that the written introduces between what is intended to be conveyed and what is actually conveyed, is typical of an originary breach that afflicts everything one might wish to keep sacrosanct, including the notion of self-presence.
>
> This originary breach that arche-writing refers to can be separated out to reveal two claims regarding spatial differing and temporal deferring. To explicate the first of these claims, Derrida's emphasis upon how writing differs from itself is simply to suggest that writing, and by extension all repetition, is split (differed) by the absence that makes it necessary. One example of this might be that we write something down because we may soon forget it, or to communicate something to someone who is not with us. According to Derrida, all writing, in order to be what it is, must be able to function in the absence of every empirically determined addressee (M 375). Derrida also considers deferral to be typical of the written and this is to reinforce that the meaning of a certain text is never present, never entirely captured by a critic's attempt to pin it down.

Here is another explanation (which I found on reddit and that greatly demystified arche-writing for me):

> Arche-writing is the principle of writing as classically determined within the tradition of metaphysics, the notion of the instituted trace that "functions" in the absence of any author or reader. As such, it is always "already there" in the sense that language, including spoken language, consistently and crucially relies on just this notion. One way to think about this is to note that whereas the voice is temporal and sequential (expressions are spoken one after another and vanish when the voice stops speaking), writing is spatial and simultaneous (all the words on a page are there at once). Derrida argues that this simultaneity is the principle of language as a whole (according to the structuralist account, which is the one Derrida engages with, for example in the famous second chapter of *On Grammatology*): the fact that a signifier, for example, can mean any one particular thing depends on there being, at the same time, other signifiers that mean other things and which are absent but constantly effective in the system. This constant "trace" of the absent signifier in the "presence" of any other signifier is the difference that creates meaning; both speech and writing as forms of language depend on it. Derrida's point with arche-writing is that if we want to identify this trace-as-fundamental-principle with one of these forms, we're going to pick writing, not speech.

It is important to know that Foucault's argument here is possibly four-fold, he argues 1) Theologically, Ecriture introduces, or substitutes for the idea of an "Author-God" of sorts, 2) Critically, this is equivalent to the notion of hidden meanings, and so on, 3) Ecriture grants a primordial status to writing, and restates a belief in some form of its origins. 4) Ecriture re-establishes a fundamental theological notion of sorts related to a "fixed and continuous tradition". (This one is in particular related to what Foucault identifies as the nature of discourse)

Now we can identify where 1) and 2) come from. "Derrida also considers deferral to be typical of the written and this is to reinforce that the meaning of a certain text is never present, never entirely captured by a critic's attempt to pin it down." implies, using transcendental arguments that there is indeed an actual "meaning" of a text (if not, then how can it be repressed?). Derrida talks about "repression" in this sense, through the notion of a signifier never referring to itself so that meaning is never truly captured, and thus indeed one can see that repression implies something originally to exist in the first place that is repressed. [5]

The argument for 3) is that one sees from the explanation copied from a reddit comment (hah why do I write it like a citation) that "the fact that a signifier, for example, can mean any one particular thing depends on there being, at the same time, other signifiers that mean other things and which are absent but constantly effective in the system." gives a primordial status to arche-writing. It identifies a

system to exist before writing even starts (the absent signifiers that are constantly effective in the system). Thus, we substitute the author-god for arche-writing, and continue the idea of an "origin" for writing.

However, there is a possible subtlety here, because arche-writing may refer to an origin for all writing. Then an origin becomes unproblematic—if all writing has an origin, this is still consistent with no unique origin for a specific work, and is equivalent in some sense to an idea of no work having an origin. Thus there is also no issue with any "creative nature" of the work. Whether arche-writing as an origin for all writing is an unquestionable claim is another point, but one that we will not engage with.

But, this is not a problem. Arche-writing establishes a system to exist in the first place, and if we take into account the tradition of examining a work through its structure, then this gives an origin for *each work* (the idea is that the system of signifiers and its working is different for each text, anyway the argument is then applied to each of these texts). So either all writing is homogenous (there is no way to obtain a different meaning from two texts that say similar things because the structure is the same across all texts) or we have an origin for each work. Presumably the former is a possibility that should be obviously untrue, or unwanted.

Finally, 4). "All writing, in order to be what it is, must be able to function in the absence of every empirically determined addressee", relates to the idea of the author surviving beyond their death which is what Foucault talked about earlier. In religious terms, the ability of writing to function in this way also reminds of us the religious tradition of how scriptures and teachings are impervious to change. Foucault's critique is through how he understands systems of knowledge and thought in terms of discourse; from the Stanford Encyclopedia of Philosophy, "archaeological method is that systems of thought and knowledge (epistemes or discursive formations, in Foucault's terminology) are governed by rules, beyond those of grammar and logic, that operate beneath the consciousness of individual subjects and define a system of conceptual possibilities that determines the boundaries of thought in a given domain and period", and therefore he rejects the idea of a fixed and continous tradition of this sort. [6]

PART II

The Author Function

If, as Foucault says, we are to liberate ourselves once and for all from this tradition, we ought to examine what happens in the wake of the disappearance of the author, and what purpose an author actually establishes in discourses.

What is the use of "Author" in discourse? Does it merely point towards an individual? One can see that this is perhaps not true, an author's name constitutes of something similar to a proper name (as talked about in the work of John Searle), it is a description of sorts. It is not equivalent of course, but

one understands that when a proper name is used in discourse, it points to more than just an individual. "When we say 'Aristotle,' we are using a word that means one or a series of definite descriptions of the type: 'the author of the Analytics,' or 'the founder of ontology,' and so forth". But they're not entirely descriptive, if a work of Aristotle were to be found to be someone else's, we cannot entirely say that the meaning of the proper name has been modified, insofar as it's not used for that particular description. A proper name constituted of both description and designation, and is not entirely determined by either.

However, the author's name is more than just the proper name. The relationship between the individual pointed out through a proper name is not the same as an author's name and what it points to. For instance if a physical characteristic of an individual were to be found to be wrong, the author's name would see no alteration, it would continue to refer to the same individual. However, if we say we found that various works of Aristotle were written by someone else, Aristotle as an author's name would indeed be modified (referring now to two individuals), whereas the proper name Aristotle would continue to refer to the very same person.

One can continue to give examples in this way, for instance how an author's name may not point to a specific individual at all (many individuals at once could be one "author", see for example the mathematical group Bourbaki who wrote various texts under the same name), but what this all gets us at is that an author's name is not used a part of speech which could be replaced by a pronoun or such, it is a function that is used in discourse for classification. When we use the name Kafka, we point to various things, among which include various works that are "Kafkaesque", or the works written by the individual "Kafka".

Discourse consisting of an author's name is also granted a special privilege in the manner that it is more than just a discourse regarding the works itself. It allows, for example, for a relationship between such works to emerge. What Foucault establishes, thus, is that an author ought to be imagined not through their creative nature or as an individual (necessarily) but as a function of the discourse, the author-function. He considers four points that characterise the discourses related to this "author-function".

First, the author-function is related to "legal-codification", a form of appropriation of works with authors, related to the penal code. Historically, the works who were assigned real authors were the ones who were often "blasphemous", and the presence of the author dictated the dangers of writing. By introducing the legal codification of such works, through copyright, Foucault claims that the bipolar nature of such discourse (between the lawful and the unlawful) has been reintroduced through the presence of the "author-function", reinstating the dangers of writing.

Second, the author-function is not universal or constant in all discourse. Historically the works that required an author were for tests of legitimacy in the scientific works, their validity was determined by their author. On the other hand literary works were widely circulated without the presence of an author (folktales, greek plays, etc.). However, a reversal happened where now scientific texts are

judged through their content and not through their author, whereas the task of traditional criticism has been to consider the author in literary works. [7] Literary anonymity is considered a puzzle to solve, and Foucault compares the task of the construction of the author that traditional criticism is concerned with to the manner in which Saint Jerome wrote about the authentication of Christian texts. [8]

Third (he never writes third in the essay), the presence of the author is used in criticism to ascribe a certain unity onto works. "The author explains the presence of certain events within a text, as well as their transformations, distortions, and their various modifications (and this through an author's biography or by reference to his particular point of view, in the analysis of his social preferences and his position within a class or by delineating his fundamental objectives). The author also constitutes a principle of unity in writing where any unevenness of production is ascribed to changes caused by evolution, maturation, or outside influence. In addition, the author serves to neutralize the contradictions that are found in a series of texts." There is a belief that there exists a certain psychological level inside the author's thought where contradictions get resolved, the incompatible elements shown to derive from a singular creative nature. Finally, the author is seen to manifest equally in all works, and drafts, letters, etc.

Fourth is the concern of how the author-function relates to a "shifter", the signs present inside the work that refer to the author—for instance the "I" in the text. With works with no author, such shifters are assumed to refer to the actual writer or an individual, but this does not happen with works with an author. For instance the "I" in a first-person novel is understood to not refer to the author. Similarly, Foucault argues, that it is as invalid to identify the writer of the work as the author as it is to identify the writer as the narrator. This phenomenon is not limited to a certain kind of literary text either, in academic texts for instance, there is a multitude of "I"'s, the "I" in the preface, the "I" who writes the main text [9], etc. In this manner, the author-function does not create a system of dependency where the other "I"'s are duplications of one, rather the author-function helps simultaneously give rise to a "plurarity of egos, and to a series of subjective positions that individuals of any class may come to obey"

But of course, this is a fairly limited conception of the author-function. Foucault also shows the complexity of the author-function by pointing to a certain kind of author, ones that can be called "initiators of discursive practices". I'll write more about this later...

§

FOOTNOTES

1. I use the word "traditional criticism" a lot in this article, so I ought to clarify what I mean by this. I use traditional criticism to substitute for the sort of criticism that is based on the quest of hidden meanings, or author's intentions that examines the creative nature of the work.

2. Ethical in this sense just means that an author *ought to* use this a style-rule of sorts. I suppose I am not a very ethical author.

3. He cites these examples as "obvious", alongside Flaubert, but to me they were not very obvious. I have tried to write down the essence of their style but here is some commentary that probably helps.

   > Flaubert's lean and precise writing style has had a large influence on 20th-century writers such as Franz Kafka and J. M. Coetzee. As Vladimir Nabokov discussed in his famous lecture series: "The greatest literary influence upon Kafka was Flaubert's. Flaubert who loathed pretty-pretty prose would have applauded Kafka's attitude towards his tool. Kafka liked to draw his terms from the language of law and science, giving them a kind of ironic precision, with no intrusion of the author's private sentiments; this was exactly Flaubert's method through which he achieved a singular poetic effect. The legacy of his work habits can best be described, therefore, as paving the way towards a slower and more introspective manner of writing."
   >
   > — *Wikipedia, "Gustave Flaubert" (https://en.wikipedia.org/w/index.php?title=Gustave_Flaubert&oldid=1377904952#Legacy), citing Nabokov,* Lectures on Literature*, vol. 1 (1980), p. 256*

   And,

   > Novelists should thank Flaubert the way poets thank spring; it all begins again with him. There really is a time before Flaubert and a time after him. Flaubert decisively established what most readers and writers think of as modern realist narration, and his influence is almost too familiar to be visible. We hardly remark of good prose that it favors the telling and brilliant detail; that it privileges a high degree of visual noticing; that it maintains an unsentimental composure and knows how to withdraw, like a good valet, from superfluous commentary; that it judges good and bad neutrally; that it seeks out the truth, even at the cost of repelling us; and that the author's fingerprints on all this are paradoxically, traceable but not visible. You can find some of this in Defoe or Austen or Balzac, but not all of it until Flaubert.
   >
   > —*James Wood,* How Fiction Works *(2008), p. 29, quoted in Wikipedia, "Gustave Flaubert" (https://en.wikipedia.org/w/index.php?title=Gustave_Flaubert&oldid=1377904952#Legacy)*

   Personally I am not exactly sold on this. I find this commentary ironic. First, this seems to be terribly biased towards realist works, and what is it to "judge good and bad neutrally"? Based on whatever academic philosophy I do know of, and some Lacanian view points that I think in, there is no "truth" that can be sought out in a neutral manner, even the act of noticing and transcribing what one has noticed is not free from the influence of ideology one lives by. At the end it doesn't matter too much, I think, because one can argue that each work is a representation of a specific reality, which itself is a function of the "author" it is related to. Maybe this does end up killing the author, because if all the work is seemingly in this reality, there is no "traces" of the author one can found (in a sense placing the author into the position of an Abrahamic god). Ironically this does reinstate the idea of an author-god. (Foucault argues this too! But I figured this out before I read what he wrote :D)

4. A little bit of my own thought, but one can also question this structural criticism of works. Are works to be considered on their own right only? The structural criticism of a work probably ought to also deal with a comparison with other works, and in this way a certain relationship exists with some special works, the works with the same writer (I avoid author here to use writer as an empirical, and not discursive property). But then this again beckons the question of what a work is—are the letters to Melania a work from Kafka? Should they be used for comparison with the other texts written by him? In this case probably not, but there are more hazy boundaries.

5. Allegedly, Derrida also denies this is the form of repression taking place, but I have not read any of his primary works, so I will withhold judgment.

6. A little aside but allegedly Derrida did anticipate many of the theological criticisms of écriture. When giving a lecture regarding this, someone apparently told him that the notion of trace is the same as the god of negative theology to which Derrida replied it is and it is not. There exist in wide circulation memes these days about "Socrates being a rage-baiter" because of the very nature of the Socratic Method, but these are spread by idiots. However, perhaps they do apply to Derrida somewhat more aptly.

7. Both this and the last point are where Foucault points out this burden placed on the author through writing, that was initially dissolved. A tangential video essay of interest here is https://youtu.be/rZwlAcQVzDE (https://youtu.be/rZwlAcQVzDE).

8. A little unfortunate that much of this is a fairly eurocentric analysis of literature, I do wonder if other things could be seen once we expand our horizons.

9. Perhaps also the "I" who writes the footnotes.