Birth of the Big Data Author?

How authorless and generative poetry paradoxically challenges the declaration of "the Death of the Author" by Roland Barthes

Andreas Refsgaard
DDK
aref@itu.dk

Abstract

The intention of this essay is to challenge and question French theorist Roland Barthes proclaimed "Death of the Author". Barthes argues that one should not treat a text as something coming from a specific person, but instead look to how it is received by the reader rather than how it is intended by an author. Although this line of thought fits well in today's culture of remix and mash-ups, a recent poetry installation made at IT University of Copenhagen made me question whether utterly disregarding the author when interpreting a text is also valid in cases where the author consist of more or something else than a single writer. By comparing and analysing my own digital installation "Start a sentence" with the book "One million stories" by Peter Adolphsen and drawing on the theories of *generative art* (Galanter 2003), *conceptual blending* (Prager 2012) and Barthes himself, the essay attempts to ask the questions of how authorless and generative poetry paradoxically challenges the declaration of "the Death of the Author".

Start a sentence…

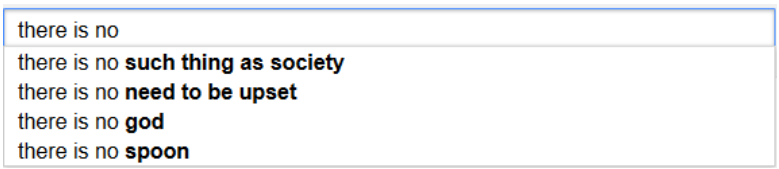


Figure 1: Screenshot of Google Suggest in action

This essay is based on a digital work of art made at IT University of Copenhagen during the fall of 2013. In collaboration with fellow students

Anders Schaumann and Morten Bang Jørgensen I built an interactive poetry installation based on Google Suggest and Microsoft Kinect as part of the course “Digital Creative Practice”. In brief, our installation, named “Start a sentence” took short speech-inputs from users and ran them through Google’s Suggest API. Up to ten sentences, based on the speech-input were projected onto a big canvas in a new visual form that enabled users to erase and rearrange the sentences to create poems or other kind of texts by using their hands as controllers via a Microsoft Kinect sensor.

The installation was coded so that all text appearing in the installation was the result of interaction between users and Google Suggest. No text was written beforehand, which made the writing of poems an activity more or less without an author per se. But what was the role of Google in this interaction, and where do the sentences Google Suggest provide come from? Before I try to answer these questions, let’s turn to the theory of Barthes.

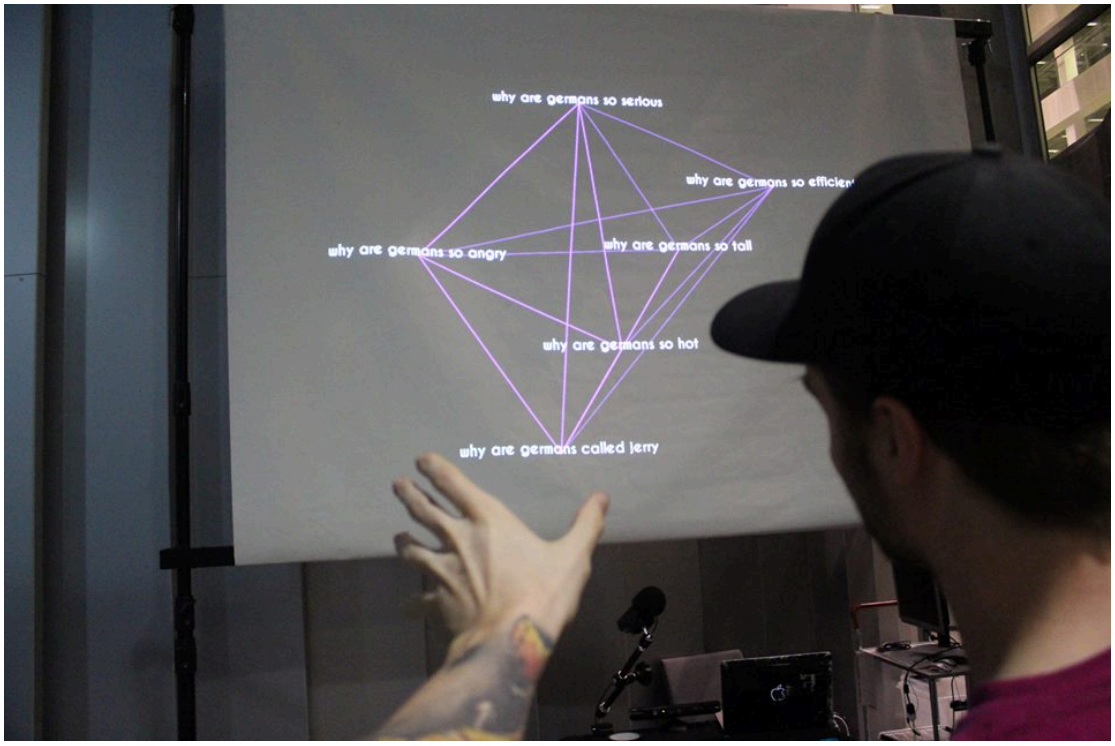


Figure 2: The “Start a sentence” installation at an exhibition at ITU, November 2013

…but who will finish it?

> The absence of the Author is not only a historical fact or an act of writing: it utterly transforms the modern text (Barthes 1998: 4)

In his essay “Death of the Author” French theorist Roland Barthes proclaims that the meaning of a text depends on how it is received rather than how it is intended. Instead of finding a "single 'theological' meaning (the 'message' of the Author-God)," critiques and scholars must realise that writing, in reality, constitutes "a multi-dimensional space," which cannot be "deciphered," only "disentangled" (Barthes 1998: 5). Barthes argues that the view of a text’s unity “lies not in its origin but in its destination” (Barthes 1998: 6), thereby shifting the focus of interpretation from author to reader, making it meaningless to force a single definitive interpretation upon the text:

> To give an Author to a text is to impose upon that text a stop clause, to furnish it with a final signification, to close the writing” (Barthes 1998: 5)

For Barthes neither the author nor the text can ever be truly original, since the text takes life according to how the reader interprets the writing as a collage of diverse sources:

> a text consists of multiple writings, issuing from several cultures and entering into dialogue with each other, into parody, into contestation; but there is one place where this multiplicity is collected, united, and this place is not the author, as we have hitherto said it was, but the reader (Barthes p. 6)

According to Barthes the author is merely a "scriptor" who still produces the text, but has no privileged position relative to the reader and does not precede or exceed the work.

It has become a measure of acknowledged notion that online everyone has the potential to simultaneously be a consumer as well as producer. If we extend the term “text” to refer to the visual arts and media at large on could

say, that we live in a culture of remix and mash-ups, where "*"writing" is no longer seen as something truly original, but as a complex act of resampling and reinterpreting material previously introduced"* (Navas 2008). In this light, Barthes theory is still of great relevance, and the works of Duchamp and pop artists such as Warhol are evident examples of artists complying the thoughts of Barthes by creating *readymades* and relying on the cultural cache of pre-existing material, as opposed to trying to create art from scratch (Navas 2008). But is disregarding the author when interpreting a text also valid in cases where the author consist of more or something else than a single writer?

By comparing "Start a sentence" with "A million stories" by Danish writer Peter Adolphsen I seek to challenge and question Barthes proclaimed "Death of the Author". Despite differences in medium and interaction both works are based on generative principles (Galanter 2003), which in diverse ways erases the traditional notion of "the Author". However, I want to illustrate how this lack of author paradoxically creates the potential for a whole new type of author, not accounted for by Barthes. The essay ultimately seeks to question how the theory of Barthes can be understood or perhaps reinterpreted when the author either explicitly surrenders control of the structure and composition of the text to the reader ("A million stories") or is replaced by interaction between reader and algorithm ("Start a sentence")?

Letting Google finish the sentences

I have already outlined the basic concept of our own installation "Start a sentence", but in the following section I seek to elaborate on Google Suggests underlying algorithm and how it affects the structure of the sentences.

As is the case with the search engine itself, Google is reluctant to specify the exact mechanism underlying Google Suggest. They do provide some insight to the public and web analysts have tried scraping Google's mechanisms through experiments. However, since the parameters of Google's algorithms

are due to continuous modification, any attempt to decipher the exact workings of the algorithm can never be more than a temporary best guess.

All that being said, there is no doubt that the absolute foundation in Google Suggest is user-generated data, or more specifically, a huge database storing peoples search queries. On average more than 5 billion daily searches are performed on Google around the world (Statisticbrain.com). Google Suggest works by attempting to complete users search queries by providing a list of the "strings" most searched for, updated in real time and based on whatever input the user writes. A number of other factors such as location, search history and trending terms also affect the algorithm and determine the order of suggestions. Finally censorship prevents (at least to some extent) hateful, pornographic, piracy-related and some other controversial queries to be suggested.

Inputting short and open starts of sentences tend to yield a higher number and a more varied set of outputs than longer and more complete sentences. E.g. "*the weather is beautiful today*" only outputs *["the weather is beautiful today in spanish"]* and *["the weather is beautiful today in french"]* while the shorter and less "finished" sentence "my inner" outputs more interesting and more poetic lines of text:

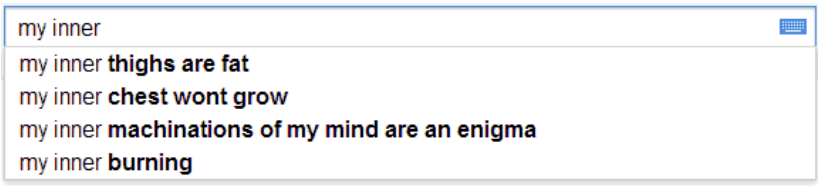
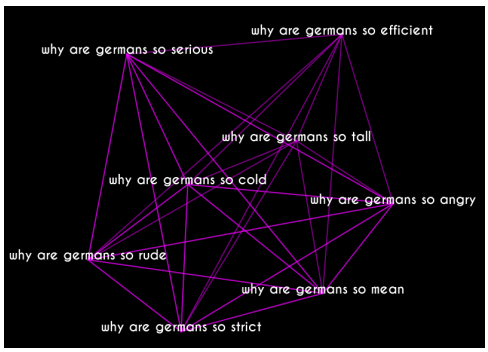


Figure 3: Screenshot of Google Suggest in action

Notice that the different lines take their beginning from the users input "*my inner*". Although in some instances this is not the case, the uniform structure of sentences establishes a stable form, nicely juxtaposing the different endings. Figure 4 and 5, both actual poems created by user while "*Start a sentence*" was exhibited at ITU, illustrates two different outcomes of this: In

"*Why are Germans so*" (Figure 4) prejudices and perceptions of German people are brought to light while "This is not" (Figure 5) flows in various directions, due to its more open input.



(Figure 4)

(Figure 5)

Because of its somewhat chaotic and unpredictable nature, an obvious tool for analysis of the poem from figure 5 is the framework the theory of conceptual blending make available:

> Conceptual blending describes the formation of figurative meanings as the integration of two or more input domains through cross-domain mappings. The mappings create a blend space in which certain input from each domain is projected into a new, emergent meaning, resulting in anything from the absurd to the poetic, humorous, or scientifically revolutionary. (Prager 2012: 2)

An attempt to map out the blend is shown in the table below:

Domain "Objects"	**Blend**	**Domain "Illusions"**
A love song A pipe* A persons hat Ikea	*Humours paradox (my interpretation):* A feeling of only recognising objects and events as being familiar, real, actually happening or just relating to the "I", when the "I" shops in Ikea (where everything is mass-produced).	Alienation Not belonging False appearance of things

* The sentence “this is not a pipe” of course refers to *“La trahison des images*” by Belgian painter René Magritte, whose statement is taken to mean that the painting itself is not a pipe, in an attempt to have the viewer question their reality. Thus Magritte's pipe stresses the binary contrast between the two domains, objects vs. illusion, which emphasizes the paradox. As such, the sentence adds depth to the poem, and poetry from Google Suggest is often filled with these peculiar meta-layers.

Though somewhat filtered and censored Google Suggest provides a unique insight into the “aggregated” human behind, the topics that mankind as a whole seems interested in, and the big and small problems it faces. The quantity of data behind the algorithm is important, not because it necessarily reflects reality 1:1 or is statistically reliable, but because, regardless of filtering, an immense amount of search queries from real people all over the world will precede each outputted sentence.

A million stories

With the book ”One million stories” (original title: ”*En million historier*”) Danish writer Peter Adolphsen follows the footsteps of Raymond Queneau, a leading figure in the movement OULIPO *(Ouvroir de littérature potentielle – “Workshop For Potential Literature”)*, whose work “*Cent Mille Milliards de Poèmes*” (“*A Hundred Thousand Billion Poems”)* he makes a modified version of. Adolphsen’s book ”One million stories” is ten pages, each page divided and cut into 6 strips, numbered 0-9. By flipping the strips the book enables the reader to choose, combine and read a million different stories. A reader can choose different principles; Let the included ten-sided die rule, read a story corresponding to a certain date (using the digits as date-date-month-month-year-year), pick favourite numbers, or simply explore the texts themselves, ignoring the numbers.

The first strip reads:

> Den gamle hanelefant harkede en klat slim op på kakkelvæggen. Den silede langsomt nedad, fulgt med øjnene af både dyret og dyrepasseren, Hektor. Begge huskede de, at

Whereby the reader may choose among 10 different sequels, for example:

> han havde udtrykt skepsis mht. kunstens forløsende kraft. “Er det ikke bare kulørte lamper på kapitalens damptromle?” lød spørgsmålet. Derpå fulgte håndgemæng og

The use of assonance (the repetition of vowel sounds e.g. the H’s in *hanelefant*, *harked*, *Hektor* and K’s in *kunstens*, *kulørte*, *kapitalens*) creates internal rhyming within the sentences of each strip, and each sentence has its own logic, lyrically and thematically, that collide with that of the next sentence in playful manners.

By forcing users to interact with the structure of the text, Adolphsen voluntarily enters into the role of Barthes’ scriptor. Adolphsen produces the words and sentences, but the work as such is "eternally written here and now" (Barthes 1998: 4) with each re-reading.

An attempt to analyse this initial part of the story with help from theory on concept blending is shown in the table below:

Domain “Zoo”	**Blend**	**Domain “Dispute over the role of art”**
Male elephant Zookeeper Spit on tile wall The observing of spit slowly running down the wall	An abstract discussion and fight over the revolutionary potential in art between a big animal in captivity and its keeper	Art as a *commodity* vs. art as *emancipatory* Critical Theory and the neo-Marxist Frankfurt School Discussion and scuffle

As Barthes would argue there can never be a 'secret, ultimate meaning' to any text, and following this line of thought the key to interpretation lies within each reader. Speaking strictly for my self, a former student of sociology, I see the story as a fable of modern capitalist society where the working class (in the shape of the caged elephant) questions whether the administrative elites (represented by the zookeeper) plans to educate and emancipate the working class through supporting the arts has any effect[1] or in fact takes focus from the structural inequalities in capitalist society. However, this is only one interpretation, and the interactive nature of the book emphasizes the range of possible texts and readings within the text, both on a literal level and a more figurative meta-sense.

Making art with generative systems

A common denominator for both "Start a sentence" and "A million stories" is the use of generative systems in the designs. Galanter defines generative art as any art practice

> (…) where the artist uses a system, such as a set of natural language rules, a computer program, a machine, or other procedural invention, which is set into motion with some degree of autonomy contributing to or resulting in a completed work of art. (Galanter 2003: 4)

Whereas "A million stories" uses the analog principles "Start a sentence" uses Google Suggest's algorithm to determine the output of users' inputs. In both cases the use of generative principles is a way of creating unexpected outcomes and combinations while at the same time leading the reader (or user) towards a high level of involvement and perhaps getting him or her to feel ownership over the produced texts.

However, this only works because the generative systems don't result in either banal or hypercomplex outputs that the readers cannot comprehend:

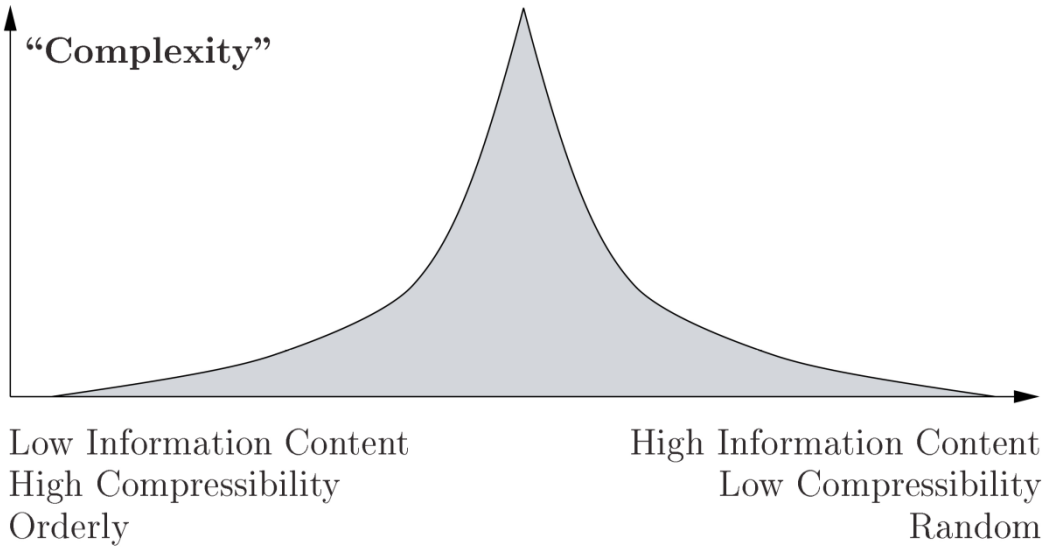
[1] In Denmark this dispute between the "rindalists" and the Social Democratic Party, trying to "educate the people for their own best", has been brilliantly described in the book "Kampen om sandhederne" by Rune Lykkeberg (2008)

> Working artists understand that an audience will quickly tire of both a highly ordered and a highly disordered aesthetic experience because both lack any structural complexity worthy of their continued attention (…) In terms of our human ability to extract meaning from a given experience we require a mix of surprise and redundancy, i.e. a signal somewhere between extreme order and disorder (Galanter 2003: 8-9)

Galanter defines the level of information in a generative system by its AIC (algorithmic information content). A computer outputting the same pattern of simple strings would have a low AIC and quickly become trivial to the human mind. A randomised output of sounds, visual forms and text, not following any specific rules would have a high AIC, but would be to overwhelming to be interesting or enjoyable for most humans. Therefore Galanter focuses more on *effective complexity* (EC) in which:

> systems that are highly ordered or disordered are given a low score, indicating simplicity, and systems that are some where in between are given a high score, indicating complexity (Galanter 2003: 10)

Figure 6 below illustrates the effective complexity as a region intermediate between total order and complete disorder.



(Figure 6)

Both “Start a sentence” and “A million stories” function as artworks because of high effective complexity. Their output is too complex to be trivial, but still not too complex or uncontrolled to derive meaning from. They operate within clear boundaries and limitations, but still provides outputs so unpredictable yet cohesive, that readers and users find them interesting to interact with and explore the limits of.

In contrast to “One million stories”, “Start a sentence” has no limit of combinations or, and during the time it was exhibited, users came up with new types of use we as designers had not foreseen or intended. Relating to EC, this enormous numeric complexity is counterbalanced by the fact that each phrase Google Suggest outputs is written by people within the rules of a language we understand. Furthermore we give the users the possibility to erase sentences they find incomprehensible, thereby letting them filter the noise, or more technically, turning down AIC to heighten the EC.

A new Big Data Author?

“A million stories” by Adolphsen is, like its source of inspiration “Cent Mille Milliards de Poèmes” by Queneau, an exciting lyrical experiment underlining the many possible readings and interpretations of a text. But by breaking the structure and stability of the text into small manipulable parts, thus giving the "power" over the text to the reader, Adolphsen confirms rather than challenges the idea of "Death of the Author": He steps out of the role of the author and into the role of the scriptor whose work is "eternally written here and now," with each re-reading, because the "origin" of meaning lies exclusively in the "language itself" and its impressions on the reader (Barthes 1998:4).

Yet by writing all the words Adolphsen is still the official author of his book (what Barthes would call a scriptor), whereas “Start a sentence” on the contrary works as an installation with no stable author and no unambiguous

scriptor. Because “Start a sentence” lets users interact with huge amounts of user created data, everyone who use Google to search online are, whether they like it or not, co-creators of the work. But are they scriptors? Authors? Regardless of their status, an interpretation of the poems from “Start a sentence” is hard to carry out while fully ignoring the fact that each sentence of every text it projected to the canvas because millions of people have searched for online. It is no longer just words, and an interpretation that only focuses on relationship between words and reader misses the point.

This is a decisive difference from the way Barthes wants us to interpret texts. The idea of an objective real interpretation of the text, or the author's original intention with it, can of course be rejected, since no single author with clear intentions exists. “Big Data texts” can be extremely interesting because they are not written by an author but by an algorithm, based on the sum of more than 5 billion linguistic actions per day. Right now, it is the closest we get to humanity as a subject, a quantitative qualified insight into the "aggregated human" wonder, problems, thoughts and dreams in all its diversity.

Conclusion

Although "Start a sentence" and "A Million Stories" has various similarities regarding the use of generative principles, the possibility of interaction e.g., only one of them truly sheds new light on Barthes theory. Where Adolphsen, by handing over the power to determine narrative and chronology from Author to reader in some way, seeks to affirm Barthes, the data-driven poetry in "Starting a sentence" makes it almost impossible not to relate to the author as a new and vital different Author figure.

Literature:

Adolphsen, Peter 2007: *En million historier*. Samleren.

Barthes, Roland 1998: "The Death of the Author" in *Art and Interpretation: An Anthology of Readings in Aesthetics and the Philosophy of Art*. Ed. Eric Dayton. Peterborough.

Galanter, Philip 2003: *Complexity Theory as a Context for Art Theory*. Retrieved from www.philipgalanter.com/downloads/ga2003_paper.pdf December 13th 2013

Lykkeberg, Rune 2007: *Kampen om sandhederne*. Informations forlag.

Navas, Eduardo 2008: *The author function in remix.* Retrieved from http://remixtheory.net/?p=309 December 13th 2013

Prager, Phillip 2012: *Making an Art of Creativity: The Cognitive Science of Duchamp and Dada* in CREATIVITY RESEARCH JOURNAL, 24(4), 1–12.

Queneau, Raymond 1961: *Cent Mille Milliards de Poèmes*. Gallimard.

Sullivan, Danny 2011: *How Google instant autocomplete suggestions work*. Retrieved from http://searchengineland.com/how-google-instant-autocomplete-suggestions-work-62592 December 13th 2013

Web references:

http://www.statisticbrain.com/google-searches/ visited December 13th 2013