

Cohesion Without Coherence: Artificial Intelligence and Narrative Form

Commentary on “I was received by the city as I stepped into the world again”

HANNES BAJOHR

Translated by Kayla Rose van Kooten

I.

¶1 The boundaries of the arts do not only exist between them, but also within these arts themselves. The following text will focus on the boundary between narration and its other—that is, between narrative and non-narrative forms. What interests me here is something that I would like to call *surface narration*: the mere form of storytelling. This, I believe, is a defining characteristic of texts created with AI.

¶2 There is every indication that AI-generated text will soon become an everyday phenomenon for readers. It is assumed that such scriptural automation will be easier to achieve in some genres than in others—especially if their artificial nature is not in question. Today, certain types of texts—from weather reports to product descriptions—are already so underdetermined that their origin seem negligible. Nevertheless, it is still generally presumed that they are written by humans (at least if one pauses to consider the question), because the technology for their sufficient automation has not yet been fully developed. This, however, is rapidly changing, and we may soon get used to attributing the origin of certain

written text to an AI. Especially with “unmarked” texts like the ones mentioned, the question of origin—whether a text is “natural” or “artificial”—may eventually become so unimportant that it will not even arise anymore. At that point, we would then be dealing with *post-artificial* texts.¹

¶3 This is not only due to their functional nature, but also to their specific structure. This includes the relative absence of style, as seen in the data-driven weather report, or the pastiche-like predictability of the hyped-up rhetoric of marketing language, which, as a function without an argument, can be adapted to promote a wide range of products. Even more significant, however, is the fact that these genres are “small” in a way that they achieve a certain balance in the relationship between *coherence* and *cohesion*.

¶4 If cohesion refers to the way text elements are linked at the phonological, orthographic, and lexico-grammatical levels, coherence pertains to

1—I discuss the difference between natural, artificial, and post-artificial texts in: Hannes Bajohr, “On Artificial and Post-Artificial Texts: Machine Learning and the Reader’s Expectations of Literary and Non-Literary Writing,” *Poetics Today* 45, no. 2 (2024): 331–61.

their abstract meaning context. The separation is ideal-typical, so that in any concrete work there is always a mixture of cohesion and coherence, as Holger Schulze explains:

¶5 “Pure coherence—as it would exist in immaterial, pure ideas—is as unthinkable as pure cohesion—a completely context-free operation with content. Every concrete artifact consists of coherence patterns that are mediated by cohesion, and cohesion patterns that become recognizable only through coherence. [...] Coherence and cohesion are inseparably linked, so that only a certain dominance of coherence or cohesion patterns can be determined.”²

¶6 In functional genres, I believe, we tend to find a roughly equal distribution of coherence and cohesion, which leads to both being more or less *unmarked*. While there are the differences in style I mentioned earlier, such cases are neither radical *cohesion texts* nor radical *coherence texts*. The former, characterized by an emphasis on structural linkage, was the subject of investigation by modernist avant-gardes—think of Henri Chopin’s sound poetry—whereas the latter leans more towards rigidly formalized arguments, such as those found in analytical philosophy. A particularly extreme example would be Richard Montague’s version of categorial grammar, which begins with the assertion “I reject the contention that an important theoretical difference exists between formal and natural languages” and proceeds to demonstrate this through cascades of formalizations like “ $\partial \in C_5 \ \& \ \varphi \in C_1 \ \& \ \langle \partial, \varphi, \psi \rangle \in R_4 \rightarrow \psi \in C_1$ ” for the disappointingly simple sentence “John loves Jane.”³

¶7 Insofar as modern natural language processing

(NLP) aligns with the tradition of such formalization attempts—albeit no longer in their technical implementation, as deep learning has shifted away from rule-based transformations to rely instead on the “distributional hypothesis”⁴ of statistical signal processing—the hope that coherence will arise naturally through cohesion runs parallel to the idea that semantics can be conjured up solely through syntax. While the latter case is not so implausible—I address it elsewhere under the heading of “dumb meaning”⁵—the question of substantive coherence naturally poses problems for text-generating AI: where coherence is subordinated to cohesion, the meaning of the linkage is always only a secondary effect of the rule that organizes its elements.

¶8 Beyond the equal distribution of cohesion and coherence in unmarked texts, as well as the extremes of pure sonority and mere logical linkage, there is something that *reflects* coherence on the level of cohesion and which is not insignificant in our cultural tradition of *marked*, namely literary texts. Montague’s operator “ \rightarrow ”—the material implication, the “if S, then P”—is elevated in these texts to the principle of organizing the material, namely

partially appropriate in that, of course, the “ideas” are still conveyed through the cohesion of logical connections. Nevertheless, it seems to me to come closer to the idea of “pure” communication of thought than natural language. Montague’s goal is to formalize the relationship between the meaning of sentences and their syntactic structure, i.e., to derive the meaning of a complex sentence from the meanings of its parts and the way they are linked together. Since the meaning of the sentence is read from the relationship between its components in their logical structure, and not from the mere sequence of words or the grammatical correctness of the sentence, as would be the case in a cohesion-heavy example, I therefore assume a tendency towards coherence here.

4—Zellig S. Harris, “Distributional Structure,” *Word* 10, no. 2–3 (1954): 146–162.

5—Hannes Bajohr, “Dumb Meaning: Machine Learning and Artificial Semantics,” *IMAGE* 37, no. 1 (2023): 58–70.

2—Holger Schulze, *Das aleatorische Spiel* (Paderborn: Fink, 2000), 23.

3—Richard Montague, “English as a Formal Language,” in: *Formal Philosophy: Selected Papers of Richard Montague*, ed. by R. H. Thomason (New Haven: Yale University Press, 1974), 196. The example is only

as “therefore.” In literary texts, “therefore” is not merely a logical consequence but also a causal one, and it becomes the greatest guarantor of cohesion: it organizes their elements into a narrative. Narrative always poses the question of a meaningful sequence of events in space and time. And this meaning is typically conveyed not by the sequence “and ... and ... and ...,” which is merely an aggregated conjunction,⁶ but by the sequence “therefore ... therefore ... therefore ...,” which is to be understood as causal in the broadest sense.

¶9 The reason for this is: correlation does not imply causation.⁷ The truism of all empirical research

6—Precisely for this reason, the conjunction has been recommended as an alternative to classical narration—most prominently in the interpretation of Deleuze and Guattari, who oppose the tree to the rhizome and, accordingly, the root-book to the book as a war machine, see Gilles Deleuze and Félix Guattari, *A Thousand Plateaus*, trans. Brian Massumi, (Minneapolis: University of Minnesota Press, 1987), 5. There, the positive invocation of the conjunction can also be found: “The tree imposes the verb ‘to be,’ but the fabric of the rhizome is the conjunction, ‘and ... and ... and ...’ This conjunction carries enough force to shake and uproot the verb ‘to be.’” (25) That ‘to be’ implies causality is not stated here, but seems plausible. In practice, particularly the *nouveau roman* made the absence of causality as a dissolution of narrativity into a structural feature, see Frank Kermode, *The Sense of an Ending. Studies in the Theory of Fiction* (Oxford: Oxford University Press, 1967), 19. Illuminating as an illustration of current popular writing practice: Juan S. Guse pointed me to a recording of writing workshop by Matt Stone and Trey Parker (the creators of *South Park*), who use the idea of causality as a prime rule in their own writing: “If the words ‘and then’ belong between those beats [what the authors call self-contained scenes], you’re fucked—you have something pretty boring. What should happen between every beat you have written down is either the word ‘therefore’ or ‘but’”—because the “but” is, too, causality, if a subverted one. “Writing Advice from Matt Stone & Trey Parker @ NYU” (2017), www.youtube.com/watch?v=vGUNqq-3jVLg (June 30, 2023).

7—Illustrative of this rule is the website *Spurious Correlations*, which collects such doubtful correla-

also applies to the limitations of narration in NLP. As computer scientist Judea Pearl tirelessly emphasizes, causality is not AI’s forte—neither in the old symbolic paradigm nor in the new sub-symbolic or statistical approaches.⁸ Computers can only process correlations, which reveal nothing about the causal coherence of their linkage. For this reason, literary scholar Angus Fletcher has expressed the belief that narration is impossible for a computer because it can only process reversible equations such as $A = B$ and $B = A$, instead of causal and irreversible relationships of the form $A \rightarrow B$.⁹ The computer could only correlate the data “fire” and “smoke,” so that “smoke, therefore fire” would appear just as plausible to the system as “fire, therefore smoke.” Correlation is not only without cause, but also without *time*. Narration, on the other hand, emerges from both and can therefore only be thought of causally.¹⁰

¶10 There are several objections to this interpreta-

tions as “Per capita consumption of mozzarella” and “Yearly engineering degrees awarded,” overlaying two causally unrelated but similar-looking statistics, see www.tylervigen.com/spurious-correlations. Another example of confusing correlation and causation are forms of magical thinking like the “cargo cult”—the alleged hope of some Pacific island tribes to attract American airplanes and their cargo by constructing symbolic runways; there is a lively debate about whether this description is based on a misunderstanding by western observers, see Ton Otto, “What Happened to Cargo Cults? Material Religions in Melanesia and the West,” *Social Analysis* 53, no. 1 (2009): 82–102.

8—See Judea Pearl and Dana MacKenzie: *The Book of Why: The New Science of Cause and Effect* (New York: Basic Books, 2018).

9—See Angus Fletcher, “Why Computers Will Never Read (or Write) Literature: A Logical Proof and a Narrative,” *Narrative* 29, no. 1 (2021): 1–28.

10—For this reason, too, one should at least mention E. M. Forster’s well-known distinction between story and plot as variants of narratives: “We have defined a story as a narrative of events arranged in their time-sequence. A plot is also a narrative of events, the emphasis falling on causality. ‘The king died and then

tion, not least that it overly reduces literature to the function of narration and problematically equates mere temporal sequences with cause-and-effect relationships. Hume had already raised an objection against such interpretations when he fundamentally questioned causality as a “necessary connexion”¹¹ (whereas Kant postulated it as the “analogy of experience” indispensable for the inner structure of knowledge).¹² Nonetheless, Fletcher’s thesis remains compelling as it offers a starting point for empirical experiments: if causes and effects as narrative engines are fundamentally unteachable to AI, one can still observe which productive errors arise in the *attempt* to do so.¹³

¶¹¹ Instead of evenly distributing coherence and cohesion or projecting one onto the other, AI texts would strive to simulate coherence *through* cohe-

the queen died,’ is a story. ‘The king died, and then the queen died of grief,’ is a plot. The time-sequence is preserved, but the sense of causality overshadows it. [...] If it is in a story we say ‘and then?’ If it is in a plot we ask ‘why?’” E. M. Forster, *Aspects of the Novel and Related Writings* (London: Arnold, 1974 [1927]), 60. Forster does not deny that time and causality are directly related, but does make a difference in where the *emphasis* on these aspects falls. In view of the temporal confusion in the Jawling text, however, it is also doubtful that language models can achieve complex regular temporal successions, as I explain below.

¶¹¹—David Hume, *An Enquiry Concerning Human Understanding*, ed. by Stephen Buckle (Cambridge: Cambridge University Press, 2007 [1748]), 59.

¶¹²—Immanuel Kant, *Critique of Pure Reason*, ed. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998 [1787]). Between Hume and Kant lies the reality of statistics, which both acknowledges that causality can never be established with absolute certainty and uses a probability calculus that is quite reliable when combined with randomized studies, see David Spiegelhalter, *The Art of Statistics: Learning From Data* (London: Pelican, 2019), chap. 4.

¶¹³—This approach is based on the assumption that errors—both glitches and bugs—are able to reveal the fractures of a system analytically, and that this access favors artistic experiments in particular, see Hannes Bajohr, *Schreiben in Distanz: Hildesheimer Poetikvorlesung* (Hildesheim: Universitätsverlag Hildesheim, 2023), 55.

sion; the result would be correlation texts in which causality only appears as a surface effect and the narration of “therefore” emerges as an almost, but never fully, successful chaining of “ands.” Just like there is “dumb meaning,” there would then be something like “dumb narrative.” Coherence effects are without a doubt partly due to attributions from the reception side—we *want* to establish a connection between the facts that are expressed between two consecutive sentences.¹⁴ However, the very possibility of accepting such connections presupposes a certain predisposition in the output itself. This effect is what I call *surface narration*.

II.

¶¹² *Kieferling* (Jawling) and *Teichenkopf* (Pondhead) emerged during my efforts to generate such a surface narrative. For this purpose, I used the language model GPT-J, an open-source alternative to GPT-3. Although relatively small, with 6 billion parameters, it can be “fine-tuned” for specific corpora. Unlike training a language model from scratch, fine-tuning involves adjusting an already pre-trained model—which already “speaks” German—to the stylistic and content-specific peculiarities of a given corpus. It would be difficult,

¶¹⁴—I was able to test this projection effect practically in my novel *Durchschnitt*, which I still created through classical coding. The mere alphabetical arrangement of average-length sentences from the corpus of Marcel Reich-Ranicki’s novel canon produced coherence effects that made unrelated elements sound like a narrative for a moment. Here is an example from Chapter “D” that creates narration solely through temporal adverbs like “then” (*dann*) and “thereafter” (*danach*): “Then she put the box back in her desk and left the key in its usual place. Thereafter, one could see the starry sky in all its purity up to the edges of the heavy clouds in the northwest. Then, however, Frau Permaneder-Buddenbrook began to cry loudly in the middle of the street and in the face of so many people.” Hannes Bajohr, *Durchschnitt* (Berlin: Frohmann, 2016), 42. According to the only person I know who has read the book in its entirety—my publisher Christiane Frohmann—reading it causes headaches.

for example, to get GPT-J to output complex and correct sentences based solely on the writings of Georg Büchner. Since language AIs model statistical dependencies over word distributions, Büchner's work would simply be too small to reach the critical mass required for sufficiently high-quality output. On the other hand, if you already use a model that has learned from a broad selection of German documents and has examples of syntactic correctness and semantic regularities, it can be steered in a more stylistically specific direction. Essentially, this process places a distinct "voice" atop the foundation of its existing German language competence.

¶13 In this case, I pursued a voice based on four contemporary novels that Elias Kreuzmair examined in an essay titled "Die Zukunft der Gegenwart (Berlin, Miami)": Berit Glanz's *Pixeltänzer* (2018), Joshua Groß's *Flexen in Miami* (2020), Julia Zange's *Realitätsgewitter* (2016), and Juan S. Guse's *Miami Punk* (2019).¹⁵ According to Kreuzmair, these works exemplify a form of writing that is not generative—neither produced by classical codes nor current AI models' neural networks—yet succeeds in describing, as "literature of the 'digital society,'" a social situation in which the digital has become commonplace. Notwithstanding their content, then, these texts are "conventional" novels; although they use postmodern self-reflexivity to consider their own standpoint in many respects, they are unquestionably classically *narrative* literature. As such, they seemed particularly attractive to me for AI training. Because, in terms of both their subject matter and their structure, they tell the story of the digital, I was curious to see what remains of both of these aspects when the texts are used as a dataset for a language model, specifically GPT-J. (I rearranged the resulting

15—Elias Kreuzmair, "Die Zukunft der Gegenwart (Berlin, Miami)," *Digitale Literatur II*, ed. Hannes Bajohr and Annette Gilbert (Munich: edition text+kritik, 2021), 35–46.

texts into a novel, which I published in the fall of 2023 under the title *(Berlin, Miami)*, drawing on Kreuzmair's essay. This work provides a broader empirical basis for the theses presented here.)¹⁶

¶14 Already the first text introduced *Kieferling* and *Teichenkopf*. This casting is not yet a narrative, but rather an example of the remarkable propensity for neologisms in large language models, which are quick to learn morphological regularities—such as the suffix "-ling" or the fact that the German language forms compound words. At the same time, it says "das Teichenkopf," using the grammatical neuter instead of "der," the correct masculine form. Interestingly, this indicates that the model has not fully internalized the rule that the gender of a compound word is determined by the gender of its "head"—after all, deep learning begins without a fixed set of rules, but extracts the "rules" from the probability distribution of the data itself.¹⁷ Nevertheless, even in the first

16—Hannes Bajohr, *(Berlin, Miami)* (Berlin: Rohstoff, 2023). In addition to GPT-J, I also employed GPT-NeoX during the writing process—a significantly larger model, though its output appeared almost indistinguishable to me. Minor discrepancies between the chapter presented here and its version in the novel stem from my editor David Frühauf's meticulous corrections. For a detailed discussion of the novel's precise stylistic choices and the rationale behind permitting minor editorial adjustments, see Hannes Bajohr, "Niemand schreibt allein: Nachwort to *(Berlin, Miami)*," *(Berlin, Miami)*, 239–273.

17—After completing the first version of this text, I was made aware that "Kieferling" is one of the many names of the Slippery Jack mushroom (which grows in symbiosis with pine trees [*Kiefernbaume*]) and that there is a 499 meter high elevation somewhere in the upper Lahn area called "Teichenkopf." I cannot say with certainty whether or not Kieferling and Teichenkopf are true neologisms—such as language models regularly produce—or the effect of a mere context shift of known tokens. The incorrect gender of "Teichenkopf" and the textual and semantic proximity of "Kiefer" (jaw) and "dentition" in the first sentence of the text at least seem to suggest that neologisms are at play.

sentence, we already encounter characters who, along with the first-person narrator, form the basis of the narration, grounding and advancing it.

¶15 The title alone—"I was received by the city as I stepped into the world again"—suggests a temporal connection that, while not yet realized causally, already carries a narrative quality. It almost evokes the prelude to a *Heimkehrergeschichte* (story of return), a hallmark of German post-war literature. Examining just the opening paragraph, one notices the contemplative tone of the first-person narrative voice: the reflection seems to take the memory of Jawling and Pondhead as a starting point for a childhood memory. This, in turn, introduces a central mystery involving a fourth character as the story's true focal point: "what had happened to my father."

¶16 The allure of this text—at least for me—lies in these characters. The Jawling appears as a kind of Odradek: difficult to visualize despite its description, yet somehow anthropomorphic—or at least possessing a semblance of subjecthood. Kafka, at least, describes Odradek as "a flat star-shaped spool for thread,"¹⁸ whereas the Jawling is merely said to be "rooted in strength," a characterization only marginally more ambiguous than that of the Pondhead, which is described as "steadied through an extremely slim base"; that there is an "unconventional division" between the two, somehow indicated by the narrator's "dentition," does not clarify matters.

¶17 However, like Odradek, Jawling and Pondhead are not just objects, but quasi-persons—animistic talismans or magical creatures—poised on the threshold between life and non-life. Holding the Jawling or the Pondhead in one's hand—it becomes evident only midway through the text that the narrator probably has both of them with

him—elicits feelings of affection, beauty, and security. Yet these emotions are undermined by the fact that the Jawling has "the task of taking down the Pondhead." The role of Jawling and Pondhead in the relationship between the father and the narrator figure remains as enigmatic as the relationship between the latter is indeterminate—at times it seems tense, as when the father takes away the Jawling, and at other times it expresses affection, such as when the paternal kiss initiates the journey to the outlook over the city.

¶18 The very existence of characters may already signal a proximity to storytelling, as their presence as actors implies narrative potential. From the outset, the introduction of these characters is closely linked to the organization of the temporal levels, such as the analepsis of childhood and the prolepsis of the father's story in the first paragraph. This linkage is so seamless that it initially obscures how fluidly and unobtrusively the temporal structure of the text keeps shifting. The phrase "in such times" in the fourth paragraph clearly indicates a different past than the one from the narrated childhood, and the "while" it takes "to reach the outskirts of this city" is again different from the pivotal moment when, about halfway through the text, the concrete deployment of the *histoire* takes place, situating it temporally as a "morning in the first week of June" and spatially in the father's house.

¶19 Perhaps it is not only the striking characters but also the abundance of remarkable phrases that distract from these inconsistencies. Stemming from a prose diction, they may mitigate the pressure for narrative coherence solely through the display of literariness. Some sentences are constructed with an almost virtuosic flair, ranging from the laconic discipline expressed in the statement, "Life in the city is not the same as life in the region," to the excess of the extravagant hypotaxis of the opening paragraph: "I suspected that the Jawling's legs had

18—Franz Kafka, "The Cares of a Family Man," *The Complete Stories*, trans. Willa and Edwin Muir (New York: Schocken Books, 1983), 469.

bitten into the indifferent Pondhead, expressing a future of community by this immediate gesture, not just in the sense of possibility—since there were already many people there—but also through the deployment of all the other means: through the rasterizations of its own layers, through an uncomplicated combination of bowels and gutters of sweat, through the counter-condensation from hundreds of valuable experiences, through the re-production of progress.”

¶20 From the path “that leads a human body shape towards suicide,” to the description “The city beneath us lay frighteningly quiet on aggravated feet,” to the point where “the feeling in the shoes had also somewhat solidified,” these formulations could compete with the more complex examples of contemporary literature. The cohesion is high, which may enhance the impression of a narrative nexus even more, as the result invites the projection of coherence. (At this point, one can begin to ask about the origin of these turns of phrase. Does the “future of community” stem from Guse? Does the “uncomplicated combination of bowels and gutters of sweat” trace back to Glanz? Yet this inquiry into origins seems increasingly tenuous when considering that all four novels exist merely as statistical dependencies within the latent space of a *single* language model.)

¶21 Nevertheless, these formulations are islands of meaning, more compelling in their self-containment than as components of a narrative chain. Consequently, what might be called a narrative flow occurs only intermittently, at least by the standards of conventionally narrated literature. The coherence markers of character constellation and temporal arrangement are opposed by a centrifugal movement, which makes it difficult to say *what* the story is that is being told here. Father and child meet, view the city—which one?—from an elevated position only accessible by stairs, while the two mysterious “heads” play an import-

ant but indeterminate role. Why this excursion? The father “wanted to show me my admiration,” “Only so that you can see the people in your happiness”—the swapping of action and being acted upon displaces characteristics and feelings onto others, frustrating any clear answer. Such confusions act as narrative glitches, much like the irritation that arises in the reading flow when the Jawling, which is firmly in the father’s hand, suddenly appears “silk-like, [...] in my [the narrator’s] fist.” As the hike concludes with the reflective sentence “Thus, our excursion was only a small part of explored feelings, a small part of the world as I knew it from my heads,” the impression of *surface*, of the shape of a story without interrogable substance, is confirmed despite all narrative effects.

*

¶22 At the same time, I hesitate to make this judgment. Would I reach the same conclusion if I didn’t know how the text came into being? After all, every text is a surface for its readers. As long as cohesion is present, coherence is a question of the tolerance for its absence, and perhaps the difference between correlation and causality matters less when only the reception decides what connection should exist between two events. Moreover, if a text aligns with literary diction, the need for such a connection may diminish further. Likewise, recurring characters—even when their actions are erratic and their motivations mysterious—could already serve a binding function in themselves: Jawling and Pondhead would possibly be inventions that already in their own right tell a story, regardless of whether it is actually told. Their presence introduces a latent potential that does not have to be executed in order to unfold on the reception side. And in the end, it is only this reception side that matters, when it is no longer clear how and in what way a text came into being, whether through AI or humans—when it has become *post-artificial*.