

How Stories Work by Manipulating What We Don't Know

Gaps as the Generators of Narrativity

The article discusses Meir Sternberg's theory of gaps, in particular regarding the fundamental linkage it establishes between gaps and narrativity. The focus is on three distinctive characteristics of Sternberg's approach. First, he conceives of gaps as referential (i.e., relating to the storyworld) – as distinct, for example, from Wolfgang Iser, for whom 'gaps' designate pretty much whatever the text has left unsaid or unconnected. Second, Sternberg makes a clear distinction between relevant and irrelevant informational lacunae – 'gaps' vs. 'blanks' or 'indeterminacies' (the latter category being largely parallel to Roman Ingarden's *Unbestimmtheitsstellen*). Finally, Sternberg recognizes three types of narrative gaps, in accordance with the three master-types of interest that, according to him, constitute narrativity: suspense, curiosity, and surprise. Of these, curiosity gaps, as being both perceptible (as opposed to surprise) and past-oriented (as opposed to suspense) tend to be perceived as prototypical, often to the detriment of the two other types, even though all three are structurally and functionally interrelated. Trying to somewhat redress this imbalance, the article concludes by illustrating the importance of considering the future-oriented suspense gaps.

1. Introduction: Sternberg's Theory of Gaps

The concept of 'gaps,' as well as the related notion of the 'gap-filling' activity of the reader, have become quite common in studies of narrative texts, especially in the context of how the reader constructs, or reconstructs, meaning from these texts. However, various uses of the term *gaps* – as well as similar or closely-related terms – show considerable conceptual disparity behind terminological similarity, regarding both definition and use in analysis. In what follows, I would like to discuss Meir Sternberg's theory of gaps and, hopefully, illuminate its distinctiveness and importance, in particular concerning the fundamental linkage it establishes between gaps and narrativity. I intend to present this theory – or at least some major aspects of it – 'synchronically' rather than diachronically, but let me just note that when looking at Sternberg's publications one may see various stages of its development over a period of about two and a half decades, beginning with the 1968 article "The King through Ironic Eyes: Biblical Narrative and the Literary Reading Process," co-authored with Menakhem Perry and originally published in Hebrew.¹ This first presentation of Sternberg's ideas about gaps is not highly theoretical, but provides a very detailed practical demonstration of how the concept of the gap can be used in analysis through a close reading of the biblical story of David, Uriah, and Bathsheva (2 Samuel 11). This story is

instrumental for dramatizing the nature of the reader's sense-making activity of gap-filling, since the text's highly laconic style (even by the standards of the Hebrew Bible's storytelling art) provides a particularly large scope for the reconstruction of implicit meanings. Later important theoretical elaborations (alongside further detailed analyses of specific narratives) appear in Sternberg's books *Expositional Modes and Temporal Ordering in Fiction* (1978, esp. 50–55, 238–246) and *The Poetics of Biblical Narrative. Ideological Literature and the Drama of Reading* (1985, esp. 235–258). Finally, the most abstract and theoretically rigorous treatment of the topic, especially regarding how gaps are essential to narrativity, appears in the 1992 article “Telling in Time (II). Chronology, Teleology, Narrativity” (esp. 508–538).

In what follows, I would like to focus on three characteristics of Sternberg's approach; the first two tend to make the ‘gaps’ more sharply and narrowly defined than is common, whereas the third tends to make their understanding more comprehensive than usual within the framework established by the first two.

2. Referentiality

The first characteristic I want to highlight is that Sternberg conceives of gaps as referential or diegetic, i.e., relating to the represented storyworld. This view, in itself, is very wide-ranging in scope, since the missing information involved in a gap may be relevant to any aspect of the storyworld – be it an event, a motive for action, a character trait, a relationship, a viewpoint, a picture of society, or even an entire reality model. However, it is still considerably more circumscribed than the view of some other approaches – something which can be made particularly clear if we compare Sternberg's conception of gaps to that of Wolfgang Iser, another influential user of the term, who helped bring it into circulation at approximately the same period as Sternberg.² Or, more precisely, *gaps* came to be the standard English equivalent in the critical discourse for Iser's German term *Leerstellen* (literally “empty spaces”), even though in the English versions of Iser's own writings it is not consistently translated this way.

At any rate, looking at what Iser variously designates as *gaps*/*Leerstellen*, it is clear that, for him, these terms may relate to any number of levels or frameworks other than that of represented reality. Let me illustrate this with some examples from his 1974 book *The Implied Reader. Patterns of Communication in Prose Fiction from Bunyan to Beckett* (going by the original use of *Leerstellen* in the German version). In a chapter on Henry Fielding's novels, Iser talks about a gap between the “familiar repertoire in the novel and one's own observation of it,” created by the ironic commentary of the narrator (1974, 33); about how the narrative of *Joseph Andrews* produces a gap by not offering a direct depiction of Lady Booby's surprised reaction when Joseph rejects her advances, thereby leaving the reader “to provide the description [himself], using the directions offered him” (38); and

about the “vacant pages” (40) – typographical space (literal “empty spaces”) between chapters. Later, in the chapter on William Thackeray’s *Vanity Fair*, Iser discusses gaps related to the unstated (but inferable) criteria for judgments that allow the reader to separate “sham attitudes” from “true [social] reality” (112). And in the chapter on James Joyce’s *Ulysses*, Iser notes how the “vacant spaces in the text” (210) increase when unrelated themes are set side by side, as well as how gaps are created between the expectations aroused by a heading and the text that follows it (214; this happens in the “Aeolus” episode, where the format of a newspaper is imitated).

Of these six examples, only the one which concerns the depiction of Lady Booby’s surprised reaction would clearly count as a ‘gap’ in Sternberg’s sense. Overall, if we try to find some common denominator for all such uses of the term by Iser, it appears to be highly abstract or general – his ‘gaps’ seem to include whatever the text has left unsaid or unconnected, thereby stimulating the reader’s activity to complete it. Or, to use Iser’s own formulations in his book *The Act of Reading. A Theory of Aesthetic Response*, any time the reader experiences “impeded ideation” as a result of the text breaking up or suspending the “connectability of textual patterns” (1978, 189).

3. Relevance vs. Irrelevance

Within the referential framework, another important characteristic of Sternberg’s conception of gaps is a clear distinction between relevant and irrelevant lacks of information – the former being recognized as gaps and the latter called ‘blanks,’ or ‘indeterminacies.’ In order to appreciate the importance of this distinction it will be useful to make another comparison – this time with the work of someone whose phenomenology of the literary text has been an important point of reference for both Sternberg and Iser, namely, Roman Ingarden, since Sternberg’s category of ‘blanks’ is very much parallel to Ingarden’s concept of *Unbestimmtheitsstellen* (translated as “spots/places of indeterminacy”).

In his book *The Literary Work of Art. An Investigation on the Borderlines of Ontology, Logic, and Theory of Literature*, in order to illustrate this notion, Ingarden gives the following example: “If a story begins with the sentence: ‘An old man was sitting at a table,’ etc., it is clear that the represented ‘table’ is indeed a ‘table’ and not, for example, a ‘chair’; but whether it is made of wood or iron, is four-legged or three-legged, etc., is left quite unsaid and therefore [...] *not determined*” (1973a, 249, italics in the original).³ And a bit later he notes that “while reading a work we are not conscious of any [...] ‘spots of indeterminacy’ in the represented objects” (251). These “spots of indeterminacy” in the world projected by a work of literature result from the necessarily incomplete or selective nature of representation through language; each object, person, event, etc., portrayed in a literary work contains a great number of them. Thus, according to Ingarden, they constitute a major characteristic of the structure of the literary work – specifi-

cally, at the level of what he calls the stratum of represented objects (or objectivities). However, they are not aesthetically functional, and when filled out – or ‘concretized’ – by readers, this tends to be done freely and arbitrarily.⁴

Sternberg also recognizes the importance of this kind of omissions, caused by the very selectiveness of representation, with respect to the literary text and its reading process – but he also contrasts it sharply with a different kind:

The world, every piece of it inexhaustible, will not lend itself to [full] coverage by discourse. Nor would discourse cover it, even if it could, because meaning and value depend on selection [...] Two conclusions follow. The first concerns the need to separate omissions or lacunae into relevancies (“gaps”) and irrelevancies (“blanks”). This is a universal of reading that no one can escape for a moment, including those who shudder at the very mention of interpretation. To make sense is to make distinctions between what was omitted for the sake of interest and what was omitted for lack of interest: between what I called, for short, gaps and blanks. Only the former demand closure, while the latter may be disregarded without loss, indeed must be disregarded to keep the narrative in focus. (1985, 236)

The need to make this distinction between gaps and blanks is deemed to be such a basic and inescapable reality of reading since the number of things about the storyworld that we might potentially wish to know is, in principle, limitless; therefore, if we don’t recognize the need to dismiss many lacks of information as irrelevant, and begin to genuinely wonder about every possible detail or aspect of the storyworld that is not explicitly provided for us by the text, we simply wouldn’t be able to go beyond the first sentence. Thus, as readers, we have to sift the gaps – which are “the result of deliberate artistic manipulation” and “compositionally, semantically, and rhetorically crucial” – from the numberless blanks, which are “irrelevant, or at best incidental” (Sternberg 1978, 311).

It is important to add, however, that this sharp theoretical distinction between relevant and irrelevant omissions or lacunae comes along with a recognition of the possible practical difficulties of making the distinction in particular cases. As Sternberg goes on to note (following the above-quoted passage): “In practice, however, the distinction turns out as problematic as it is inescapable, since it can appeal to no formal (and thus automatically applicable) marker,” which means that “one reader’s gap may prove another’s blank” (1985, 236). One can easily think of many omitted details that most (if not all) readers would agree are too trivial⁵ to be interesting or relevant to the understanding of the narrative – for example, the color of the messenger’s eyes in the David and Uriah story (236; perhaps picking up on Ingarden’s example mentioned in note 3); and, on the other hand (or the other end of the scale), it is also possible to think of lacks of information that are obviously interesting and relevant – for example, the question concerning the identity of the murderer in a classic detective story, where the whole plot at the level of the storyworld revolves around the attempt to answer it. But there are also many cases that may fall somewhere in-between these clear extremes, where different readers or readings would differ over whether a certain lack of information should be treated as relevant or irrelevant. Moreover, in circles of academic or ‘professional’ readers, turning previously ignored blanks to gaps that deserve attention is a common strategy for producing new readings (sometimes at the level of whole schools or orientations – think,

in this context, of psychoanalytic criticism). At any rate, these issues essentially belong to the realm of interpretation, and as such raise the usual kinds of problems – as well as inviting the usual kinds of solutions – typical of interpretative questions.⁶

4. Narrativity and Three Types of Gaps

Whereas the two characteristics of Sternberg's approach discussed so far – referentiality and the distinction between relevant gaps and irrelevant blanks/indeeterminacies – tend to make his conception of 'gaps' more sharply and narrowly defined compared to various other approaches, I believe the next one I want to discuss leads to a more comprehensive understanding of gaps than usual, while highlighting their chronological and dynamic aspect. This requires me to briefly present Sternberg's conception of narrativity (i.e., what constitutes the essence, or distinctive characteristic, of narrative). Unlike the majority of narratological approaches, which define narrative in the mimetic terms of represented or narrated *action*, Sternberg defines it in the rhetorical-functionalist terms of *interest*.⁷ This narrative interest is aroused in the reader through informational gaps – that is, a manipulation of our lack of knowledge about the storyworld, regarding any aspect of this world. Let us recall a previous quote, in the context of distinguishing gaps from indeterminacies, about gaps as "what was omitted *for the sake of interest*" (my emphasis): this is, in essence, how gaps become the fundamental generator of narrativity.

Sternberg differentiates between three master types of narrative interest.⁸ All of them are based on the interplay between narrative's two temporal sequences: the mimetic, or actional, and the textual; but this interplay systematically varies, according to three fundamental types of dynamics – and, respectively, three basic types of gaps. These can be described according to a combination of two criteria: (1) Does the information withheld from the reader, via gaps, belong to the narrative past or future? (2) Is the reader aware that information is being withheld? In other words, are we aware of the existence of the gaps? Such an awareness would create an expectation for receiving the missing information – and this would result either in *suspense*, if the expectation relates to the narrative future, or in *curiosity*, if it bears on the narrative past. However, if the reader is unaware that an informational gap exists – that is, if the gap is camouflaged, or 'unmarked' – it will become perceptible and operative only when unexpectedly disclosed, thereby creating *surprise*. A slightly different way of putting it is that all three types of narrative interest, and gaps, are based on us – as readers – not knowing something about the storyworld, but whereas in the case of suspense and curiosity (which may be regarded as two subtypes of 'expectation') we know that we don't know, in the case of surprise we don't know that we don't know.

* * *

Of these three types, curiosity gaps tend to be perceived as the prototypical kind of gap, often to the detriment of the two other types; a good indication of that is how critics who don't think – at least not consciously – in terms of the three types, choose almost invariably curiosity gaps as examples. And the reason is probably that curiosity gaps feature both awareness of the missing information (that is, a marked presence of the absence) and the actual existence of this missing information within the storyworld, as belonging to the narrative past, whereas both suspense and surprise gaps lack one of these attributes. Surprise gaps, as noted, are based on lack of awareness about their existence until the surprise actually lends, when the gap, in relation to what was mistakenly considered a fully revealed past, is both revealed and filled in at the same time;⁹ whereas suspense gaps seem less 'real' because they relate to expectations about the narrative future – that is, what hasn't happened yet within the storyworld, and may never actually happen. (Of course, if we are dealing with a fictional text, one might claim that nothing is 'real' there, but I am considering how we view things *within* the parameters of the fictional world, where the asymmetry between our perception of past and future, in terms of their degree of reality, generally holds in parallel with real life.) Thus, overlooking or neglecting these two types – suspense and surprise – when discussing gaps considerably limits our ability to understand how narratives work, since all three types are structurally and functionally interrelated.

In this context, I would like to elaborate on the importance of recognizing the future-oriented suspense gaps,¹⁰ since their supposed 'unreality' in terms of what actually occurs in the storyworld highlights the productivity of thinking about narrativity from what I called a rhetorical rather than mimetic perspective. And since this unreality is particularly striking in the case of permanent suspense gaps, let me first introduce another important distinction made by Sternberg – between *temporary* and *permanent* gaps, as well as some of its theoretical consequences.

A temporary gap is one that a work opens at some point along the text continuum only to fill it in explicitly and satisfactorily – or at least to enable the reader to do so with certainty – at a subsequent stage. A permanent gap, on the other hand, gives rise to questions to which no single, fully explicit and authoritative answer is provided by the text from beginning to end. Importantly, the force of this distinction hinges to a very large extent on the text's terminal point – since prior to it, we cannot be sure to what category a gap would turn out to belong: any still-open gap may be closed, and any gap that seems closed may be reopened.

If the ending determines whether a gap turns out to be temporary or permanent, we may also reverse perspective and note that the temporary or permanent nature of the gaps crucially influences how the ending is perceived, particularly in terms of closure – namely, to what extent it is 'closed' or 'open.' And here I would like to briefly summarize some points made in my own work on narrative closure (see Segal 2010, esp.155–163; Segal 2019, esp. the sections "Ending and Closure," "Narrative Closure and Open[ended-]ness").

As a guiding principle for understanding narrative closure, we might begin with the following statement (from Barbara Herrnstein Smith's classic study of closure in poetry) about the basic relationship between a text's overall structure and its closure: "The description of a poem's structure, then, becomes the answer to the question, 'What keeps it going?' [...] It allows the possibility of a corollary question, namely, 'What stops it from going?' and immediately suggests the close relationship between poetic structure and closure" (Smith 1968, 4). If we apply this principle to narrative structure, then according to Sternberg's conception of narrativity it is narrative interest, in its three master types, that keeps the narrative going; and therefore, what stops it from going is *the cessation, or termination, of narrative interest*. Accordingly, the ending of a narrative text would produce an effect of closure inasmuch as it brings to a halt the operation of all kinds of narrative interest by filling in all of the significant informational gaps about the storyworld that have arisen along the textual sequence; in other words, these gaps turn out to have been temporary. And conversely, narrative 'openness' results from significant gaps relating to the storyworld that remain open – or at least not definitely closed – even at the end of the text; in other words, from permanent gaps. These can be both curiosity gaps, relating to the narrative past, and suspense gaps, relating to the narrative future.¹¹

Strictly speaking, what I have just defined is not the (non-)closure of narrative but rather that of narrativity. That is, I do not claim that an explanation in terms of (non-)termination of narrative interest will exhaustively elucidate everything related to (non-)closure within a narrative text. For example, the text may use all sorts of linguistic/stylistic devices, which in themselves are not specific to narrative. I do claim, however, that where narrativity is dominant in a text, an explanation of closure in terms of the play of narrative interest via the three types of gaps will always be crucially important, and that the understanding of how any other elements in the text operate in the context of closure will depend on their interaction with these workings of narrative interest.

* * *

With this theoretical framework in place, we can now return to discussing the importance of recognizing suspense gaps. And let me begin with a very basic example – a pair of sentences proposed for consideration by Gerald Prince: "The cat sat on the mat" (a favorite of grammar books) vs. "The cat sat on the dog's mat" (1995, 80). It is probably easy to agree that the second sentence features a higher degree of narrativity than the first, but what is the reason for that? According to Prince himself, it results from the presence of "some kind of conflict" (80). However, if we pause to think about it, no *actual* conflict is reported here; judged solely by its direct mimetic content, the second sentence is just as 'static' as the first, reporting on a cat that is seating on a mat. The heightened narrativity of the cat-on-the-dog's-mat scenario resides rather in its clear *potential* for conflict, which focuses our attention on possible future consequences, thus generating a suspense gap. Let me emphasize that I am talking here in relative rather

than absolute terms – about *degrees* of narrativity; the first sentence doesn't *prevent* us from thinking about possible future developments relating to the cat, but the second directs our attention to such possible developments in a considerably more structured and 'urgent' way, because of that suggestion of possible conflict.

My next example is taken from Wolf Schmid's book *The Nonnarrated* (2023), which I have chosen both because it is a very recent study (thus illustrating the relevance of the topic under discussion) and because there is a high degree of overlap between Schmid's concept of the 'nonnarrated' and Sternberg's 'gaps.' In accordance with what I have said earlier about the tendency to favor curiosity over the two other types of narrative interest when discussing gaps (or closely related concepts), most of Schmid's book – including its best analyses (e.g., of Pushkin's "The Shot" or Dostoevsky's *The Brothers Karamazov*) – indeed features cases of curiosity gaps. There is, however, a chapter on Anton Chekhov's open-ended stories (41–48) where the focus shifts to issues of suspense; Schmid himself doesn't use this term, nor makes (at least explicitly) the kind of distinction I am making here, and I believe that thinking about the issues he raises in terms of suspense gaps can clarify the picture.

My attention was particularly drawn to this discussion of Chekhov by the chapter's opening sentence: "A common reason for which something is not narrated is that the text ends earlier than the story being told" (41), followed a bit later by: "Often the uncertainty [in the open endings of Chekhov's stories] is due to the fact that the narrative ends earlier than the story it tells" (41).¹² Which raises an interesting question: How can something be considered part of "the story being told" if it is not actually narrated, especially if this non-narrated part belongs to the story's 'future'? Note that this is *not* a case where something that is not explicitly stated by the text can be safely inferred, since the whole point of the open-endedness Schmid is talking about is that we genuinely don't know what the future of the storyworld holds.

Of Chekhov's stories discussed in this context, I would like to look specifically at "The Teacher of Literature," where the titular character, Nikitin, seems at first to happily achieve the fulfilment of his dream – a union with the woman he loves – but then goes through a bitter process of disillusionment regarding both the nature of his supposed marital happiness and, more generally, the vulgar and suffocating environment in which he lives. The story ends with the following words which Nikitin writes in his diary:

My God, where am I?! I am surrounded by vulgarity and more vulgarity. Boring, insipid people, pots of sour cream, jugs of milk, cockroaches, stupid women... There is nothing more terrible, more offensive, more stultifying than vulgarity. I've got to get out of here! I've got to get away this very day, or else I will lose my mind! (Cathy Popkin's translation, qtd. in Schmid 2023, 42)

However, notes Schmid, "It is not told whether this entry in the diary exhausts the hero's whole departure or whether he actually takes practical steps to change his life" (42) – and this is where the open-endedness of the story lies.

So, according to Schmid, the possibility of Nikitin taking practical steps to change his life, as well as the contrasting possibility, where he doesn't go beyond

noting his feelings in the diary, are part of “the story” – only the text, as he puts it, ends before the end of the story. But in what way can both these possibilities be said to be part of the story? Well, like the “conflict” between cat and dog in relation to the sentence I quoted earlier from Prince, the various possible developments in Nikitin’s future can be “part of” Chekhov’s story only in terms of its *structure of narrative interest*, namely, as part of a permanent suspense gap;¹³ the rhetorical conception of narrativity as based on narrative interest is just as necessary here. And going back to the distinction between relevancies and irrelevancies, one might add that in this respect the above-mentioned possibilities in Nikitin’s future are unlike many (in fact, innumerable) other possible or hypothetical future developments in this storyworld that we are not told about, developments which are not part of the same structure of interest, and therefore remain indeterminacies, rather than becoming “part of the story” as gaps.

5. Conclusion

In this article, I have discussed Meir Sternberg’s theory of gaps, in an attempt to illuminate its distinctiveness and importance. I have presented two of its characteristics by comparison with the approaches of other well-known theoreticians in the field, Wolfgang Iser and Roman Ingarden. Sternberg conceives of gaps as referential (i.e., relating to the storyworld), as distinct from Iser’s much more general and vaguer conception; and he makes a clear distinction between relevant and irrelevant informational lacunae – ‘gaps’ vs. ‘blanks’ or ‘indeterminacies.’ The latter category, while largely parallel to Ingarden’s *Unbestimmtheitsstellen*, is differently understood and radically recontextualized within Sternberg’s approach. I have then turned my focus to the fundamental linkage Sternberg establishes between gaps and narrativity (the latter understood as based on narrative interest) and his recognition of three types of narrative gaps, in accordance with the three master-types of interest that constitute narrativity: suspense, curiosity, and surprise. Of these, curiosity gaps, as being both perceptible (as opposed to surprise) and past-oriented (as opposed to suspense) tend to be perceived as prototypical, often to the detriment of the two other types, even though all three are structurally and functionally interrelated. Trying to somewhat redress this imbalance, I have illustrated the importance of considering the future-oriented suspense gaps – and in particular permanent suspense gaps, which are striking in their lack of actual existence in pure mimetic terms, while still crucial to a rhetorical understanding of how narratives work.

In general, since the study of dynamic gapping is the study of narrativity itself, its importance to narratology should be clear; however, let me also briefly suggest several specific directions for further research. First, in my presentation of Sternberg’s ideas about gaps and narrative at large, I have naturally focused on constants and universals. Yet, there remains the endless variety of individual texts that can be analyzed in terms of their structure of narrative interest. Fur-

thermore, systematic variations according to authorial, school- and period-style, or generic features could richly repay further study in the context of literary history and descriptive poetics. For example, my own previous work analyzes how various subgenres of detective fiction differ according to shifting hierarchies between curiosity and suspense, as well as the relative importance of the final surprise twist (Segal 2010, 163–189).

Second, the study of dynamic gapping can focus on various textual levels. In my comments on how closure relates to the manipulations of narrative interest, I have emphasized the text's overall structure, with special attention to its final unit (whichever way it is segmented). But it is also possible to focus on smaller, micro-sequential units. For example, one might look at dialogues as producing a new curiosity gap with each turn – which may be further subdivided – in terms of what caused the character to utter it (intentions, motivations, mental state). Such gaps may vary enormously in terms of their relative importance or the difficulty of filling them in (from the most 'obvious,' which doesn't consciously register as a gap at all, to the opaquest and most enigmatic), but precisely because of that it would be beneficial to deautomatize how we perceive and process them through such a method.

Finally, while my emphasis throughout has been on literature, Sternberg's conception of narrativity is, in fact, transmedial. Accordingly, it might be productive in the context of transmedial (or intermedial) narratology to examine the similarities and differences between how various narrative media tend to produce and manipulate gaps. To mention just a couple of prominent points of comparison between literature and other media – consider the consequences of drama's much heavier reliance on dialogue at the expense of narration, or cinema's tendency toward greater descriptive fullness of externals alongside the systematic gapping of internals.

Bibliography

- Brewer, William F. / Lichtenstein, Edward H. (1981): "Event Schemas, Story Schemas, and Story Grammars." In: John Long / Alan D. Baddeley (eds.), *Attention and Performance*. Vol. 9. Hillsdale, NJ, pp. 363–379.
- Brewer, William F. / Lichtenstein, Edward H. (1982): "Stories Are to Entertain. A Structural-Affect Theory of Stories." In: *Journal of Pragmatics* 6 (No. 5–6), pp. 473–486.
- de Sherbinin, Julie W. (1993): "Life beyond Text. The Nature of Illusion in 'The Teacher of Literature.'" In: Robert Louis Jackson (ed.), *Reading Chekhov's Text*. Evanston, IL, pp. 115–126.
- Ingarden, Roman (1973a): *The Literary Work of Art. An Investigation on the Borderlines of Ontology, Logic, and Theory of Literature*. Trans. by George G. Grabowicz. Evanston, IL.
- Ingarden, Roman (1973b): *The Cognition of the Literary Work of Art*. Trans. by Ruth Ann Crowley / Kenneth R. Olsen. Evanston, IL.
- Iser, Wolfgang (1974): *The Implied Reader. Patterns of Communication in Prose Fiction from Bunyan to Beckett*. Baltimore, MD.
- Iser, Wolfgang (1978): *The Act of Reading. A Theory of Aesthetic Response*. Baltimore, MD.
- McHale, Brian / Segal, Eyal (2015): "Small World. The Tel Aviv School of Poetics and Semiotics." In: Marina Grishakova / Silvi Salupere (eds.), *Theoretical Schools and Circles in the Twentieth-Century Humanities: Literary Theory, History, Philosophy*. New York, NY, pp. 196–215.

- Prince, Gerald (1995): "On Narratology. Criteria, Corpus, Context." In *Narrative* 3 (No. 1), pp. 73–84.
- Schmid, Wolf (2023): *The Nonnarrated*. Berlin.
- Segal, Eyal (2010): "Closure in Detective Fiction." In: *Poetics Today* 31 (No. 2), pp. 153–215.
- Segal, Eyal (2019): "Beginnings and Endings." In: *Oxford Research Encyclopedia of Literature*. DOI: <https://doi.org/10.1093/acrefore/9780190201098.013.1051>.
- Smith, Barbara Herrnstein (1968): *Poetic Closure. A Study of How Poems End*. Chicago, IL.
- Sternberg, Meir (1978): *Expositional Modes and Temporal Ordering in Fiction*. Baltimore, MD.
- Sternberg, Meir (1985): *The Poetics of Biblical Narrative. Ideological Literature and the Drama of Reading*. Bloomington, IN.
- Sternberg, Meir (1992): "Telling in Time (II). Chronology, Teleology, Narrativity." In: *Poetics Today* 13 (No. 3). pp. 463–541.
- Sternberg, Meir (2010): "Narrativity. From Objectivist to Functional Paradigm." In: *Poetics Today* 31 (No. 3). pp. 507–659.

Eyal Segal

Independent Scholar

E-mail: eyalsegal3@gmail.com

How to cite this article:

Segal, Eyal: "How Stories Work by Manipulating What We Don't Know. Gaps as the Generators of Narrativity." In: *DIEGESIS. Interdisciplinary E-Journal for Narrative Research / Interdisziplinäres E-Journal für Erzählforschung* 15.1 (2026). 164–176.

DOI: [10.25926/6dwq-9702](https://doi.org/10.25926/6dwq-9702)

URN: [urn:nbn:de:hbz:468-20261008-121104-9](https://nbn-resolving.org/urn:nbn:de:hbz:468-20261008-121104-9)

URL: <https://ejournals.bib.uni-wuppertal.de/diegesis/de/article/download/1789/3023/5995>



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹ A slightly revised English version of this article has appeared as chapter 6 of Sternberg's book *The Poetics of Biblical Narrative. Ideological Literature and the Drama of Reading*, "Gaps, Ambiguity and the Reading Process" (1985, 186–229).

² In this context, the work of both Sternberg and Iser, as representatives of the Tel Aviv and the Konstanz schools, respectively, may be viewed as part of the international wave of reader-oriented criticism, which emerged during the late 1960s and early 1970s (see McHale / Segal 2015, 197–200).

³ In *The Cognition of the Literary Work of Art*, Ingarden uses as a parallel example the unmentioned eye color of a certain character (1973b, 50).

⁴ This description of Ingarden's approach involves some simplification, especially if we take into account his later work, *The Cognition of the Literary Work of Art*. (*The Literary Work of Art* was originally published in 1931 in German; *The Cognition of the Literary Work of Art* in 1937 in Polish, prior to the revised and enlarged German version of 1968.) There we find, for example, the claim that not all "permissible" concretizations are equally desirable and that "different concretizations are thus not of equal value, especially since various ways of filling out places of indeterminacy can introduce new aesthetically valuable qualities" (1973b, 54); or certain reflections on how

places of indeterminacy may be employed in a work “in an intelligible and artistically planned selection and arrangement in order to make possible an aesthetically valuable concretization” (244–245); or a consideration of how certain lacking pieces of information make us aware that we “do not actually understand everything” (290), rather than being limited to trivial material details, as in the above-mentioned examples of indeterminacies. Overall, such passages give the impression that Ingarden is groping toward an understanding of ‘gaps’ in Sternberg’s sense, but without the necessary conceptual framework to support it.

⁵ It is important to differentiate this sense of ‘trivial,’ which relates to the unimportance of the information involved, from another possible sense, which has to do with gaps that don’t require conscious and laborious consideration but can rather be filled in more-or-less automatically; in this case, the ‘triviality’ stems from the ease of the inferential activity involved. During the 8th ENN conference at the University of Wuppertal, where the paper on which this article is based was presented, I heard several uses of the adjective *trivial* in both senses, during the various sessions in the panel on gaps; one that stuck in my mind was by Marco Caracciolo, regarding the gap “how many heads does Jane Eyre have?” This is clearly trivial in the second sense but *not* in the first – in fact, the obvious ease of filling in this gap (which most readers would not even consciously register as such) might be said to go hand in hand with its importance with respect to the text’s underlying reality model – think of how the presence of a character with more (or less) than one head would affect an otherwise realist novel. It is also worth mentioning that Ingarden makes a similar distinction between the “places of indeterminacy” discussed above and “unambiguous inference[s]” or “tacit information” (1973b, 241–244).

⁶ In *The Poetics of Biblical Narrative*, Sternberg discusses some general interpretative pointers to the identification of informational lacunae as gaps (1985, 240–258). They fall into two basic types: explicit questions formulated in the text (whether by the narrator or by a character), and some perceptible fracture or incongruity which demands explanation (a juxtaposition between two pieces of reality that fail to harmonize, or a violation of recognized norms).

⁷ He calls the first kind of approaches “objectivist,” because they focus on the represented object. Sternberg (2010) is his most extensive discussion of this topic, surveying twelve types of objectivist definitions of narrative.

⁸ This trio of basic narrative forces was first presented as such by Sternberg in *Expositional Modes and Temporal Ordering in Fiction* (1978; though further systematized up to 1992, as noted above). It was picked up shortly afterward – but without the rhetorical orientation – by cognitive-empirical research (Brewer / Lichtenstein 1981; 1982).

⁹ To be more precise, when a gap emerges in what was earlier considered a fully revealed past, there are two basic possibilities: the gap itself may be filled in at the same time, but such a gap may also remain open (or at least not unambiguously closed) after its existence is revealed, thereby producing curiosity in turn. Think, for example, of a surprise produced by the sudden discovery that a dark secret lies in a character’s past: this may occur during the revelation of the secret itself or leave the exact nature of that secret unknown.

¹⁰ Interestingly, Sternberg himself seems to have had some problems with fully conceptualizing suspense in terms of gaps, since he sometimes talks about gaps as if they were limited to curiosity and surprise. I am thinking especially of recurring statements throughout both *Expositional Modes and Temporal Ordering in Fiction* and *The Poetics of Biblical Narrative* about gaps as always necessitating deformation of chronological order – e.g., “All gaps result from discontinuities between the order of narration and the order of occurrence, with its straight chronology” (Sternberg 1985, 237). This, however, applies only to curiosity and surprise gaps, since they bear on the past and in order to “direct (misdirect, redirect) notice to the narrative past, already complete at the moment of narration, you must gap and deform it out of time sequence” (Sternberg 1992, 527). But – as Sternberg himself goes on to explain – in order to produce suspense, although deformation of chronological order is an optional tool (as in anticipations of events-to-come), it is not a necessary one, since “the future contrasts with the past (the focus of the two other interests) in being naturally opaque [...] to achieve the same [i.e., directing notice] with regard to the narrative future [...] nothing need be done beyond following developments in their proper sequence” (527).

¹¹ But not surprise gaps – since in terms of interest, openness entails the reader’s awareness that some relevant information is lacking, and therefore narrative openness cannot manifest itself in a definitionally hidden surprise gap. However, it can certainly be caused by it or follow it, if we are dealing with surprises of the second type described in note 9.

¹² In fact, this description applies to the endings of only two of the three Chekhov stories analyzed by Schmid: “The Teacher of Literature” and “The Bride.” While discussing the third, “The Lady with the Little Dog,” Schmid focuses on what is a curiosity – rather than suspense – gap,

one that has to do with whether a certain inner transformation of Gurov *has already taken place* by the end-point of the story.

¹³ In discussing Schmid's reading of Chekhov's story, my focus is on a theoretical point rather than interpretation, and so I have assumed Schmid's interpretation as a given. Let me note, however, that it is certainly possible to read the story's ending as considerably less open, viewing Nikitin as undergoing a crucial internal realization while attaching less importance to the question of what external act would follow such a realization. This is how Julie W. de Sherbinin, for example, views the ending: "*Nikitin's feat is not awaiting him in the future but has already been achieved by the story's end.* His act of self-liberation moves him beyond the 'scripts' offered by the Russian literary and religio-cultural traditions. Overcoming the past, freeing oneself from the fetters of perception dictated by cultural images and literary texts, is, in Chekhovian terms, a critical juncture of self-discovery, a crucial moment of triumph for the individual" (1993, 125, my emphasis; elsewhere in the same article, what Nikitin goes through is called an "epiphany").

A more general issue that may be involved here is modernist fiction's greater emphasis on mental processes and interiority compared to earlier fiction – with Chekhov widely considered a foundational figure of early modernism. Thus, present-day readers who are already 'educated' by modernist norms can more easily view an internal development undergone by a character as providing closure, even when such a development is not affirmed by external action; on the other hand, open-endedness – also valorized by modernist aesthetics – may be perceived in itself as particularly satisfactory (in contrast to the negative critical responses of Chekhov's contemporaries, cited by Schmid [2023, 42], to the perceived open-endedness of the story).