THE ROUTLEDGE COMPANION TO PHILOSOPHY OF LITERATURE

Edited by
Noël Carroll and John Gibson

Taylor and Francis
Not for distribution
Neuroscience and Literature

William P. Seeley

Rather than treating language as information to analyze syntactically and semantically and then store in memory, language is now seen as a set of processing instructions for how to construct a mental representation of the described situation.

Zwaan and Radvansky, 1998, p. 162

The growing general interest in understanding how neuroscience can contribute to explanations of our understanding and appreciation of art has been slow to find its way to philosophy of literature. Of course this is not to say that neuroscience has not had any influence on current theories about our engagement, understanding, and appreciation of literary works. Colin Martindale developed a scientific approach to literature in his book The Clockwork Muse (1990). His prototype-preference theory drew heavily on early artificial neural network research. Jenefer Robinson has appealed to results from affective neuroscience in her discussion of our emotional engagement with literary works (see Robinson, 2005). More recently literary critics have tried to extend cognitive literary theory in the direction of neuroscience (Ashton, 2005; Byatt, 2006; Young, 2010). This line of research traces its origin to commentary on a 1995 paper by Herbert Simon (1995) in the Stanford Humanities Review. So, although there hasn’t been much going on that could justifiably be called a neuroscience of literature, discussions of neuroscientific research have woven their way into the fabric of some aspects of philosophy of literature and literary theory. Nonetheless, there has been a great deal of resistance to the creeping insurgence of neuroscience into both fields. I will challenge this skepticism in what follows. In Section 1, I will sketch a model for research at the crossroads of philosophy and neuroscience of art and suggest how it can be applied to the philosophy of literature. In Section 2, I will introduce situation models and discuss the role they have been hypothesized plays in discourse comprehension and narrative understanding outside of literature. In Section 3, I will explore some ways that situation models can contribute to discussions of narrative understanding, appreciation, and the artistically salient features that differentiate literature as a category of art from non-art literary forms like newspaper columns, magazine articles, or biographies.

1 Neuroscience of Art and Literature

Simon argued that computational models in cognitive science are ideally suited for explaining how texts evoke meanings in literary practice. There is a sense in which he
is right. Cognitive science is an interdisciplinary field dedicated to the study of the ways organisms acquire, represent, manipulate, and use information embedded in their local environment in the production of behavior. Artworks, including works of literature, are communicative devices. They are stimuli intentionally designed to trigger, induce, or evoke perceptual, affective, and cognitive responses sufficient to enable listeners, viewers, spectators, and readers to recover their content from information carried in their surface structure. This suggests that questions about our understanding and appreciation of artworks in a broad range of media can be fruitfully treated as questions about the ways consumers acquire, represent, manipulate, and use information carried in their formal-compositional structure in order to recognize and evaluate their content. Neuroscience is a tool within the broad umbrella of cognitive science that can be used to investigate, model and explain how these computational processes are structured and implemented. Therefore, neuroscience can be used to investigate, model, and explain the psychological processes that contribute to our understanding and appreciation of artworks. If successful, the outcomes of this project can potentially contribute novel information to discussions within philosophy of art.

A general model for neuroscience of art emerges from a discussion of artists’ productive practices in the fine arts. We are bombarded by a dense flux of sensory information from the environment every moment of every day. However, only a fraction of this information is germane to our current goals and behaviors. Selectivity is, therefore, a critical feature of cognitive systems. Research suggests that we solve the problem of selectivity by focusing attention on small sets of diagnostic features, or features that are sufficient to enable us to categorize, and thereby perceptually recognize, the identities, shapes, locations, and affordances of objects and events. Artists have developed a range of formal-compositional devices for culling key diagnostic features from ordinary perceptual experience and rendering them in a medium, e.g. the use of maquettes and formal studies to develop material in sculpture or the expressive power of gestural lines in caricature. An analogous story can be told about literary texts. Semantic comprehension and recall track the gist of a story, not the lexical-semantic structure of the sentences that make up the explicit surface structure of the text. The argument, then, is that the formal-compositional strategies employed in fine art and literature work as communicative strategies because they are directed at sets of features sufficient to evoke a mental representation of some desired content via the ordinary operations of cognitive systems. This, in turn, entails a tight coupling between the formal-compositional strategies employed by artists and authors and the operations of cognitive systems that can be used to explain our engagement with artworks.

The strength of this model is that it appeals to no more psychological resources than those needed to explain ordinary cognitive and affective behaviors. But this is also a shortcoming. One is still owed a story that links explanations of how artworks work as perceptual, affective, and cognitive stimuli to the ways we recognize the artistic salience of their formal, expressive, and semantic features. Skeptics take this conundrum as a general stumbling block for any psychologically grounded theory of art (Tallis, 2008). There is something to this worry. If psychological explanations of our engagement with artworks are on par with psychological explanations of their non-art cousins, if we need to appeal to something over and above the resources of psychology and neuroscience to differentiate the two, then the work of understanding art may already have been done in figuring out how to sort them. However, this objection rests on a misunderstanding of research methods in psychology and neuroscience. Theories of art are just that, theories.
At the very least experimental research in neuroscience of art is germane here because it can be used to do just what experimental research is designed for: where competing theories yield clear hypotheses and predictions, clear experiments can be designed to evaluate alternative hypotheses.

There is also a stronger position one can adopt. Consider the productive practices of naturalist landscape painters. There is no single ideal solution to the problem of realistic depiction, e.g. Rackstraw Downes, Robert Bechtle, Thomas Cole, John Constable, and Jacob van Ruisdael all adopted different formal-compositional strategies as solutions to this problem, conventions for pictorial depiction that partially define the artistic salience of their works. Painters, therefore, have to choose how to execute their subject matter. The same can be said of literature. There is no one ideal way to tell a story or construct a narrative. Authors must choose the perspective from which to construct their treatment of the subject matter of a play, short story, novel, or poem—e.g. Sophocles, Euripides, and Jean Anouilh’s tellings of the story of Antigone. The critical constraints on this aspect of artists and authors’ productive practices are the nature of the medium, the range of conventions and practices that define artistic salience within a given community, and the artistically salient aesthetic, expressive, semantic, and thematic content he or she intends to convey. This entails that the formal-compositional strategies employed to articulate the content of a work naturally carry information diagnostics for its artistic salience.

There are three core questions that frame the development of this model in literature: what are the formal-compositional strategies that define the works of a particular author, literary school, or historical period as categories of art? How are these strategies used to carry and convey diagnostic information? What accounts for the artistic salience of these strategies and works? Of course, answers to these questions rarely sort out cleanly in contemporary contexts. Works of literature often straddle different literary and textual categories. Gabriel Garcia Marquez’s novel *The General in His Labyrinth* is ambiguously located on a continuum between literary fiction and historical biography. Likewise, his journalistic essay *The Story of a Shipwrecked Sailor* lies ambiguously between an historical record and a literary novel. Nonetheless, our engagement with artworks is guided by knowledge of the productive and evaluative conventions that shaped its content, a fact that may be of more importance in contemporary contexts where it is the norm for works to overlap multiple genres or reside in interstitial spaces between categories of art. These three questions, therefore, frame a heuristic strategy that can help determine how artists and authors have employed diagnostic cues and guide the development of research programs in neuroscience of literature.

2 Narrative Comprehension: A Case Study

Suzanne Langer (1953) defined the productive practices of different artforms relative to what she called their *primary illusion*, what we might alternatively call their *essential abstraction*. I am skeptical about the medium-specific, essentialist tone of Langer’s view. However, the concept of an essential abstraction can be used as a heuristic for understanding research in neuroscience of art. The abstractness of artistic media is the source of their communicative power. Viewers have to recover the rich, three-dimensional depictive content of naturalistic paintings from their abstract two-dimensional surfaces. Likewise, listeners have to recover the expressive content of works of pure music from temporal aspects of their tonal, harmonic, and rhythmic contours. In many cases these
cognitive processes involve the mediating influence of world knowledge, art historical knowledge, and knowledge of an appropriate range of art critical appreciative conventions. These cognitive factors shape not only how we approach a work of fine art, but also what we recognize in it and how we perceive or experience it. The question then for a neuroscience of literature is, how do authors exploit the abstractness of their medium to communicate the contents of their works?

2.1 Situation Models and Narrative Comprehension

A situation model is an elaboration or articulation of the actions and events described in a literary text. Situation models emerge from the integration of information explicit in the surface structure of the text and a reader’s knowledge base (Zwaan, 2004). The net result is a multi-dimensional model of the referential situation described in the text—a mental representation that tracks changes in the presence and locations of characters, their goals, intentions, and actions, the presence and locations of objects, causal relations among characters and objects, and the general spatio-temporal structure of the situation described. Sentence level semantic comprehension can be used to illustrate the explanatory power of situation models (Zwaan, 2004):

The ranger saw an eagle in the sky.
The ranger saw an eagle in the nest.
The midfielder scored a goal.

A flying eagle has outstretched wings for soaring and flexibly, adapting to changes in speed, elevation, wind, and the locations and movements of targets and obstacles. The wings of a nesting eagle are pulled in close, folded against its body as it lies still resting in its cramped quarters. This dynamic spatial information is readily available to readers who know what an eagle is, even though it is not explicitly presented in the text. Likewise, readers familiar with the game of soccer are quickly able to reconstruct the event described, including a general assessment of the players involved and their rough spatial relationship with one another. These examples demonstrate that, even at the level of a single sentence, what is articulated in a situation model is vastly richer than what is explicitly present in the text. This is the abstraction that drives literary practice. Of course, there is also great deal of ambiguity in each of these cases. Is the eagle in the sky soaring, diving, or flapping its wings flying away with a fish? Was the goal scored with a foot or a head, by a kick or a lunge, from outside the penalty area or in close? We’re not sure. At this stage these situation models are openly structured. They present a broad range of narrative possibilities waiting to be pruned as more information becomes available.

The ambiguities associated with the sentences discussed in the previous paragraph point to a key structural feature of situation models in narrative contexts. The content of a narrative is not rigidly determined in a bottom-up fashion by the order of presentation of information. Narrative comprehension is more like the recurrent ebb and flow of tides than the linear tumble of a river racing to the sea. It is a tentative exploratory behavior whereby readers collect novel bits of information that retroactively articulate the goals, intentions, causes, and broader thematic connections that that structure local narrative events and integrate them into a cohesive whole. The development of situation models is an analogously symmetrical process. Consider this set of sentences (Zwaan, 2004):

The ranger saw an eagle in the sky.
The ranger saw an eagle in the nest.
The midfielder scored a goal.
The red squirrel jumped from the oak to the pine tree. The red fire truck came swerving around the corner.

The initial encounter with the word ‘red’ is semantically underdetermined in both cases. Its semantic content is retroactively articulated by the occurrence of the noun it modifies. Just as the word ‘red’ articulates the semantic content of the following word, e.g. a smaller European squirrel with tufted ears as opposed to its larger eared North American cousin, it is retroactively articulated by what follows it, e.g. the brownish red color of the squirrel vs. the cadmium red of the fire truck. Analogously, at a larger scale, situation models are constructed, updated, revisited, and revised as the broad range of narrative possibilities each initially suggests is retroactively pruned by the unfolding structure of the story (Carroll, 2001).

The forward and backward articulation of situation models reflects the structure of a common and powerful narrative strategy: erotetic narration. Events described at a local scale foreground contextually salient information, and thereby suggest questions about the outcomes of possible future narrative events. These processes generate narrative expectations that readers interpret as just or unjust, that readers hope will come true or hope will be thwarted by some future set of narrative events (see Carroll and Seeley, 2013). Narrative appreciation is a measure of the degree to which narrative outcomes match moral expectations that are set up by the treatment of the story and the conventions of literary practice within a genre. For instance, we may fantasize that Robert Jordan should ride off into the mountains with Maria at the close of Hemingway’s For Whom the Bell Tolls. But this Hollywood ending would ultimately be dissatisfying because it would violate existentialist conventions that define the thematic content of the work. Erotetic narratives are, therefore, built from nested sets of existing and prospective situation models that define expectations about the moral value of the outcomes of potential narrative events at local and global scales relative to general world knowledge (including knowledge of the influence of ordinary beliefs, desires, and emotions on human behaviors) and domain-specific knowledge (including knowledge of the history of artistic production, literary genre categories, and associated evaluative conventions). On this account literary texts can be conceptualized as sparsely structured scripts that scaffold the construction of a story, recipes for assembling mental representations of the subject matter and thematic content of the narrative. Situation models therefore provide a potential psychological mechanism for elaborating the content of a literary work from diagnostic semantic cues contained in the surface structure of a text.

2.2 Neuroscience and An Embodied Account of Narrative Understanding

Situation models are embodied multi-modal representations that depend on the same range of neurophysiological processes that support the perception and performance of objects, events, and actions in ordinary behavioral contexts (Speer et al., 2009). For instance, the same somatotopically mapped premotor areas involved in motor planning, preparation, executing an action, and simulating it in kinesthetic imagery are involved in the semantic comprehension of words and sentences that describe that action. These premotor areas are part of a cortico-spinal circuit that connects premotor areas with both the muscle groups that would be used to accomplish that action and related somatotopically organized proprioceptive and somatosensory perceptual areas (Kandel et al., 2006). This entails that a broad range of somatosensory processes, including in
some cases low-level electromyographic muscle activation ordinarily associated with motor preparation and kinesthetic imagery, are involved in the semantic comprehension of action words and sentences (Candidi et al., 2010). Similarly, the same regions of the occipital and inferior temporal visual areas that are used to perceptually recognize different classes of objects are involved in the semantic comprehension of object names (Martin, 2007). Further, readers’ eye movements track the spatial structure of described scenes, actions, and events, their grip-scale tracks the size and shape of the objects that characters interact with, and adopting an appropriate grip size and shape primes semantic comprehension, e.g. the sensibility of the phrase ‘throwing a dart’ is judged more quickly when participants adopt a dart-throwing hand position (see Fischer and Zwaan, 2006). Finally, the same brain regions involved in our capacity to recognize and respond emotionally to ordinary actions and events underwrite our capacity to recognize the emotional states depicted in pictures and narrative text (Foroni and Semin, 2013).

These results generalize from the semantic comprehension of words, clauses, and behaviors to narrative comprehension in text and film (Speer et al., 2009; Zacks et al., 2009). Speer and her colleagues used functional magnetic resonance (fMRI) imagery to track changes in brain activation along the usual dimensions associated with narrative change in literary texts: spatial change, object change, character change, causal change, and goal change. They observed significant correlations between perceived change in the goals and actions of characters and areas of posterior superior temporal cortex associated with the perception biological motion in goal-directed intentional action. They also found correlations between changes in the goals of a character and areas of prefrontal cortex associated with the recognition and interpretation of the order and structure of goal-directed actions (BA9, 44, 46). Character–object interactions were associated with increased activation in the premotor somatosensory hand areas. Changes in the spatial locations of objects and characters were correlated with increased activation in frontal eye-fields and parahippocampal cortex. Frontal eye-fields are associated with eye-movements in endogenously, or internally, cued attention shifts. The Parahippocampal cortex has been associated with representations of the appearance and layout of scenes (Mullaly and Macquire, 2011). These results replicate similar findings associated with event perception film narratives (Zacks et al., 2009) and everyday perceptual contexts (Kurby and Zacks, 2007). The broad range of behavioral and neuroscientific evidence reviewed here suggests that readers actively engage perceptual, affective, and motor processes employed in ordinary behavioral and perceptual contexts to construct embodied representations of critical aspects of depicted actions and events in narrative comprehension.

How does this story about narrative comprehension fit to the general model for neuroscience of art introduced above? Situation models are a cognitive mechanism used to perceive, track, and interpret change in the local environment across a range of dimensions that define the dynamic structure of behaviorally salient events. The suggestion is that we perceptually understand ordinary events in everyday contexts by generating crossmodal simulations, or situation models, that include representations of expected perceptual change associated with our own movements, the actions and goals of others, and the general spatio-temporal causal structure of the world. These forward models enable us to predict the dynamic structure of events and actions and smoothly interact with our environment. These same cognitive processes are a strong constraint on artistic production in narrative contexts in film and literature. Consequently, a close coupling between the formal-compositional strategies used by authors and the psychological
mechanisms that underwrite the production and use of situation models in ordinary contexts can be used to explain our engagement with narrative literary works.

3 Situation Models and the Philosophy of Literature

Narrative comprehension is not unique to literature. Structured narrative connections are an efficient way to organize and deliver information we want other people to be able to comprehend and use. Stories play a critical role in many aspects of life, from the parables and oral histories we use to pass on folk wisdom to at least some of the journalistic stories we read in newspapers and magazines to keep abreast of current events. So, a key question is, “What if anything, does this story tell us about literature as a form of art?” This kind of general worry for naturalistic discussions of art often centers around a definitional problem. What if anything can neuroscience contribute to our understanding of those features unique to an artform that would enable us to differentiate exemplars of that category from ordinary objects, events, and artifacts, e.g. non-literary narrative forms like historical accounts, biographies, and parables.

Minimally, there seems to be agreement that literary works involve a unique treatment or articulation of a subject matter that is unified around some humanly interesting thematic content—the creative imposition of form on the subject matter of a work with the goal of expressing some unified, artistically salient thematic content (Lamarque, 2009; Stecker, 1996). These two aspects of literary practice define an institutional account of literature: a text can be identified as a literary work by recognizing that it was produced with the intention that it be read within the framework of some set of conventions defining literary practice for a particular literary genre, e.g. romance novels, murder mysteries, or epics and quest narratives like The Odyssey or Cervantes’ Don Quixote. Domain-specific knowledge of these conventions, in turn, defines expectations that constrain and shape the modes of attention we bring to bear on our engagement with the work. Literary appreciation is a matter of evaluating the fit of the composition and content of a literary work to the conventions governing literary practices within a literary genre. Within this context, our engagement with narrative works can be loosely divided into three, hierarchically related practices directed at recovering the meaning and thematic content of the work: explication, elucidation, and interpretation (Beardsley, 1981; Lamarque, 2009). Explication is a matter of discerning the semantic content of the sentences in a text. However, the interest is not in the literal meaning of sentences per se, but rather in their symbolic meaning, in the way their content is inflected in the context of the work as a whole, and the contribution it subsequently makes to that content. Lamarque argues that, in this regard, explication is analogous to archeology. Its task is the recovery of hidden or lost information from observations of the functionality of a sentence in the larger context of its use in a work. This claim is analogous with the claim that sentence meaning is best construed at a situation-model level as opposed to a lexical-semantic level.

Elucidation is defined as the articulation of content at the subject level, or the search for narrative details about the world of the work that are not made explicit in the text. Elucidation is thereby explicitly linked with situation models. It is the process by which situation models are prospectively and retroactively articulated to accommodate new information. It is also a process whereby more global interpretation level facts about the thematic devices employed to unify the text into a cohesive literary narrative are assimilated into situation models that articulate local narrative events. The role these devices play is derived from domain-specific knowledge of genre categories, interpretive practices,
and the art historical context of a work. The fact that situation models are articulated by domain-specific knowledge of productive and evaluative conventions entails that they naturally carry information about the range of variables that account for the artistic salience of a literary work.

Lamarque describes four interpretations of Ophelia’s madness in *Hamlet* that can be used to illustrate this process. Elizabethans saw her behavior as a consequence of female love-melancholy or erotomania. Victorian-era readers interpreted it as a consequence of sexual intimidation. Freudians saw it as guilt over a desire that her lover free her from an unresolved Oedipal attachment to her father. Feminist critics from the 1970s read it as an act of protest and rebellion against a patriarchal order of gender oppression and sexual intimidation. Situation models constructed to articulate the subject level content of the play will track the psychological causes of described events differently in each of these contexts. The task of constructing situation models of the subject of the text is therefore similarly an archeological task shaped and constrained by interpretation-level facts that define the overarching, artistically salient, thematic content of the work. We are not interested in the subject matter of the play per se, in the events and actions constitutive of Ophelia’s behavior on and offstage in the drama. Rather what is of interest is the way our understanding of their causes and effects is inflected by the artistically salient thematic content of the work. These contextual facts about the way interpretive themes articulate the process of elucidation, in turn, influence the sentence meaning and the symbolism we attribute to Ophelia’s utterances and descriptions of her behavior in explication.

Situation model theories of narrative comprehension provide a clear account of narrative understanding—the process of getting clear what the story is, of stitching its parts together into a globally cohesive whole that yields moral expectations about events and the fates of characters, of understanding how the form of the narrative has been used to articulate the humanly interesting theme that binds it together into a cohesive whole. However, it is often argued that this kind of causal-psychological account of our engagement with artworks falls short in explaining appreciative practices. Here the claim is that the causal-psychological processes involved in narrative comprehension are equally involved in our understanding of works that are done poorly as those that are done well, or to content that is fit to a literary category as to content that isn’t. The critical piece of the puzzle is, therefore, the claim that situation models naturally carry information about the artistic salience of literary works.

The artistic salience of a literary work generally lies in the way the treatment of its subject matter is used to articulate a novel or unique perspective on a humanly interesting theme. Authors accomplish this goal by shaping the subject matter of the work in such a way that the discrete episodes of the narrative thematically cohere into a formally unified whole. How does this work? The surface structure of a literary text, in the context of general world knowledge, domain specific knowledge of the genre conventions governing the appropriate literary form, and cultural knowledge of the historical context of the work, carries diagnostic cues sufficient to construct a situation model constitutive of its broader artistic content. Knowledge of this range of conventions is, in turn, a constitutive constraint on a reader’s experience of content of the work. This entails that, at least in some contexts, narrative appreciation involves something like an embodied appraisal. If the degree of fit with genre-specified conventions is appropriate, the work will be experienced as narratively and interpretively compelling. If not, the situation model constitutive of a reader’s comprehension of the narrative will be experienced as wooden, unbelievable, stylistically inappropriate and so on; for example, the stylized
plots of adolescent scifi-fantasy novels, a Hemingway novel with a happy ending, or the last summer comedy you walked out of in the movies.

This model suggests several general sources of narrative appreciation (see Lamarque, 2009; Stecker, 1996). At the most basic level there is simply a question of how naturally compelling the story is, a question of whether the goals, motives, actions, and events described stitch together in a satisfactorily compelling way that is naturally fit to our moral expectations about actions and events in ordinary contexts. At a more abstract level, as mentioned above, the treatment of the subject matter should conform to the range of formal-compositional and evaluative conventions that define literary genres, e.g. canonically literary works should display a compelling degree of realism in their depictions of character relationships that is missing from the formulaic film-noire style of true crime novels and cyberpunk science fiction. There is also a question of the cognitive value of the work, a question of whether what readers learn from the novel perspective presented in the work is fit to its genre category. For instance, the situation models constitutive of our understanding of an overly maudlin romantic comedy will be experienced as wooden and unbelievable, diminishing the cognitive value of the work, and likewise true crime novels are ill-suited to funny light romantic lessons. Finally, the archeological sleuthing involved in explication, elucidation, and interpretation are often experienced by readers as the deployment of their own creative imagination. Authors, like any other artists, employ some degree of slip between the contents of their works and the conventions defining categories of art within their media as a productive tool, e.g. as a means to promote novel perspectives and challenge the imaginative capacities of readers. The discover of some slip between genre conventions and narrative expectations is thereby often experienced as a satisfying challenge to resolve. However, these disruptive effects are a matter of degree. Too much slip can alternatively interfere with the smooth deployment of creative imagination in the construction of situation models constitutive of the content of the work, causing it to be dissatisfyingly experienced as unbelievable, clunky, or over-intellectualized.

Conclusions

What’s the take-home point in all of this for philosophy of literature? Situation models of text comprehension provide leverage in understanding how explication, elucidation, and interpretation function to help us recognize and appreciate aesthetic, cognitive, and interpretive artistic salience in literary works. Of course, this won’t explain why we appreciate these works. It won’t explain the source of their artistic value. The answer to this question likely lies somewhere in an exploration of how norms and conventions constrain evaluative judgments, just as skeptics about neuroscience of art suggest. Nonetheless, attention to facts about the way the thematic content of a work emerges in the author’s treatment of the subject matter is a valid source of artistic appreciation for literary works. Situation models provide a mechanism for modeling how the treatment of the subject matter of a work carries information about the moral expectations and artistic conventions that are the source of its artistic salience. Imaging studies and surface electromyography research measuring muscle responses have been profitably used to investigate, model, and explain how these cognitive processes contribute to narrative understanding, and as a consequence narrative appreciation, at a neurophysiological level. These studies therefore demonstrate the utility of neuroscientific research to philosophy of literature.
WILLIAM P. SEELEY

References


