

Theory of Mind and the Unreliable Narrator

LIU Xiao-yan

Nanjing University, Nanjing, China

Theory of Mind was used in the analysis of characters' mind, which provides a new perspective to study the unreliable narrator. This paper takes for its object of analysis the interrelationship between the narrator's mind and the interwoven network of the minds in the unreliable narration. The main articulations of this study are intended to reflect some of the theoretical concerns central to the unreliable narrator from the aspects of the minds of author, characters, and readers to present a full examination of cognitive mechanical formation of unreliable narration, the reality of the unreliable text world and its cognitive effort for the readers. Findings suggest that the unreliable narrator is the ones the author (implied author) chooses to represent the real cognitive world of real life and the unreliable narration question and probe the limits of human minds and human expectations, and force the reader to rethink and reconsider reactions that seem "natural".

Keywords: Theory of Mind, the unreliable narrator, the minds of unreliable narration

Introduction

Theory of Mind provides a new perspective to probe the complexity of the minds of unreliable narration. Booth made the first definition of the unreliable narrator in his famous book *The Rhetoric of Fiction* in 1961, which intrigued the wide range study of it in literature criticism. Until Phelan (2005), who was well known as the leader of postclassical rhetorical theory of narrative, inherited his guider Booth's hypothesis of the distance between the narrator and the implied author as the yardstick, extends "unreliable narration" from two axes to three axes, and illustrated its six subtypes in detail. However, he paid no attention to the cognitive aspect of the unreliable narrator and implied author. Nünning (1997) gave evidence that narrative unreliability can be reconceptualized in the context of frame theory and of readers' cognitive strategies, and claims that:

These include both textual data and the reader's preexisting conceptual knowledge of the world. In sum whether a narrator is called unreliable or not does not depend on the distance between the norms and values of the narrator and those of the implied author but between the distance that separates the narrator's view of the world from the reader's world-model and standards of normality. (pp. 83-105)

Nünning (1997) put too much emphasis on the reader's role in the perception of reliability, and neglected the work of the minds of the unreliable narrator and the interwoven network of the minds in the text world with a cognitive theory of unreliability.

The Minds of the Unreliable Narration

The unreliable narration questions and probes the limits of human minds and human expectations, and forces the reader to rethink and reconsider reactions that seem “natural”. The authors often:

Construct their characters as human beings with ToM skills similar to the reader. However, writers may also undermine readers’ and characters’ assumptions that they correctly perceive others’ intentions, by deliberately depicting mistakes in their construction of other people’s minds—or misconstruction—thereby creating complex webs of misunderstanding and misinterpretation between people of different cultures or even those growing up in the same home (as in Shakespeare’s *Othello* or *A Midsummer Night’s Dream*). (Marron, 2011, p. 187)

Abnormal minds provide a perfect laboratory for examining how human characters contend with, adapt to, and overcome—or fail to overcome—obstacles to their own survival and success. Such world allows for “extreme case analysis” and provides readers with optional cognitive “playgrounds” for their own thinking and feeling brains.

The unreliability of narrator exists not in the distance between the implied author and the narrator, or the distance between the reader and the fact presented in the text, but the distance between the mind of narrator and the other minds in the text world or his/her own minds. Oatley (2011) claimed: “A narrative is also a kind of simulation, one that runs on minds. It binds a complex of intentions, outcomes, and emotions together into a comprehensible whole” (p. 21). It is for such matters that we need not just a Theory of Mind, but Theory of Minds to study the unreliability of the narrator. According to the study of Palmer (2011):

An important part of the social mind is our capacity for intermental thought, which is joint, group, shared, or collective thought, as opposed to intramental, or individual or private thought. It is also known as socially distributed, situated, or extended cognition, and also as intersubjectivity. (p. 28)

So when we study the reliability of a narrator, we should consider not only the intramental thought, which will be varied as the unfolding of the narrator’s telling, but also the intermental thought, which is a crucially important component of fictional narrative, because:

Much of the mental functioning that occurs in novels is done by large organizations, small groups, work colleagues, friends, families, couples, and other intermental units. It could plausibly be argued that a large amount of the subject matter of novels is the formation, development, and breakdown of these intermental systems. (Palmer, 2011, p. 28)

These social minds are woven into the fabric of unreliable narration.

To some extent, the unreliable narration can be the supplement of the unnatural narration, that is, some of the unreliable narrator has the mental problem, or lack the cognitive capability of mind-reading. As the cognitive turn of literary criticism, Theory of Mind became the most important methods of the analysis of fictional text. Theory of Mind is central to read and understand literature, which is what enables us to “put ourselves in another’s shoes”, that is:

The mechanism we use to understand what is going on in other people’s heads. How we react to one another socially is the most important aspect of our lives. Without an understanding of what people think, what they want and what they believe about the world, it is impossible to operate in any society. Theory of Mind is the name given to this understanding of others. It is the basic necessity of humanity and is understood the same way the world over. (O’Connell, 1997, p. 2)

This means that we can understand, define, and describe people on the basis of their perceived (or understood) beliefs, desires, feelings, values, experiences, and intentions. So when we read a novel, we treat characters as if they are real people, and we ascribe to them a ToM. As for the reader of novels, “the only way in which the reader can understand a novel is by trying to follow the workings of characters’ minds and thereby by attributing states of minds to them” (Palmer, 2011, p. 29). However, it is difficult for the readers to construct mental frames for the unreliable narrators. Because a narrator with a mental disability or a skewed perspective is indicative of unreliability as well as the under-developed perspective of a child narrator, and who lack the basic cognitive capability for mind-reading. So according to the relationships between the unreliability of narrators and their capability of mind-reading, there are five basic types of cognitive unreliable narrator. First, the mental disordered narrators, who cannot understand of other’s thoughts, and just give the fragmented description of what they saw, such as the characters in Faulkner’s novel *The Sound and the Fury* (2009). Second, the narrators are the strong-minded people with the negative emotion, attitude, and worldview about others, which may hinder their judgment of others’ minds, such as the unreliable narrator in *Atonement* (2003). Third, their cognitive capability is underdevelopment, such as the narrator in Joyce Carol Oates’s novel *My Sister, My Love* (2008). In this novel, Skyler’s narration about the story of his murdered sister is unreliable. On the one hand, his cognitive condition made him not have the capability to explain what had happened to his sister at the night of the murder, when he was nine years old, and his sister was six years old; on the other hand, he could not make up from the trauma of that event and was in the mental hospital since that time. Fourth, can we say we really know our own thought, or that of others? In fact, we are not sure of that, especially, when we recollect the memory of past event. It involves many factors, consisted with the lost memory fragment and mystery of the past fact. At the same time, neither can we shed off the influence of the present emotion, worldview, and value, nor have a know-all ability to read other’s real intention. It’s particularly difficult for us to guess the people mind whose mind frame is beyond our existent ones. Fifth, some characters tend to be difficult to read. They may be impassivity and impenetrability. As Theory of Mind also relates to dispositions that persist over time and that form part of another’s character or personality, some personality of the unreliable narrator also inclined to be considered as unreliable. Above five kinds of minds represent most of the types of unreliable narrators, but not all of that.

“Plays, films, short stories, and novels are about people with intentions who interact with each other. People are good at understanding processes on step at a time, but less good at understanding intersections of such processes” (Oatley, 2011, p. 21). This seems to suggest humans should recognize the limitations of their normal human intuition and, when confronted with seemingly alien modes of behavior, they should adopt a more formal logical approach. Like their fictional protagonists, readers should observe others, collect empirical data and extrapolate new world—or cultural—rules. This approach can encourage readers to adopt a more flexible, and perhaps sympathetic, approach to different modes of thinking. Many writers of unreliable narration believed that their narratives could thus encourage a more pluralistic society.

Realities of Unreliable Narration

The unreliable narrator is the one the author (implied author) chooses to represent the real cognitive world of real life. “So, being able to know others and oneself requires imagination” (Oatley, 2011, p. 13). That is, what we know about our own mind and others is not so reliable. In real life, although most of us have the capability to read

other's mind, we cannot know the real mind of other's people, especially for an intruder to our circumstances. Palmer (2011) claimed that:

In cognitive terms, nearly all of our life is spent on the surface, on the outside, on the sense that all of the minds with which you are involved (with the admittedly rather important exception of your own) are only ever experienced on the surface, and from the outside. (p. 29)

In this sense, the true mystery of narration is not the reliable narration, but the unreliable narration. For most of the unreliable narrator never tell you about the real truth of the story, instead, they just told you about what they saw as an outsider. Just as the visible or appearance is the most appreciated aspect, it is not surprising that Wilde (2005) said that: "It is only shallow people who do not judge by appearances. The true mystery of the world is the visible, not the invisible" (p. 23). From this point, the unreliable narrator is representation of the reality of the real life.

There is a paradoxical assumption: We perceive people's observable behavior as both a highly informative and at the same time quite unreliable source of information about their minds.

In fact, it might be difficult for us to appreciate just how much mind reading takes place on a level inaccessible to our consciousness. For it seems that while our perceptual systems eagerly register the information about people's bodies and their facial expressions, these systems do not necessarily make all that information available to us for our conscious interpretation. (Zunshine, 2011, p. 65)

This double perspective is fundamental and inescapable, and it informs all of our social life and cultural representations. As for the fictional minds, "works of fiction magnify and vivify various points on the continuum of our imperfect mutual knowledge: spectacular feats and failures of mind reading are the hinges on which many a fictional plot turns" (Zunshine, 2011, p. 69). So the unreliable narration perfectly reflects of this kind of unreadable of mind.

We use our Theory of Mind in natural, nonconscious ways when dealing with other people, but this is not adequate for understanding the cognitive processes of nonstandard human minds, we must more consciously be aware of the types of logical, deductive steps one must go through in our mind reading process of the unreliable narration.

Even though we know that there must be a mental state behind a behavior, we don't really know what that state is. That is, there is always a possibility that something else is going on behind even the most seemingly transparent behavior. (Zunshine, 2011, p. 69)

According to the study of cognitive psychology, there would be a motivation behind all the actions. So in real world, it is easy to see others' action, gesture, facing express, and hear their words, and then we can deduct what is their mind. However, most of the time we have the wrong guess, although we can easy to guess someone is thirsty, when he pours himself a glass of water. As for more complex mind, when we want to know whether the other love you or not, you may guess from the expression, body language, and the words, but all these can lead you to the opposite conclusion. On the other hand, if you are the narrator, as the lack of adequate information or having some other consideration, you may make up some details of the story, or hide some real fact, which is normal in real life. According to Phelan's (2005) opinion, the narrator who violates the normal values of ethics is the unreliable one. But we cannot say the narrator who distracts from the normal way is the

real unreliable narrator, he may be the one who has the different opinion about the world. In this sense, at least, he is reliable narrator of his own mind. To some extent, the unreliable narrator may be the real honest narrator of the reality of social life.

So unreliable narration does not mean that there should be a reducing of reality of the text world, on the contrary, it is the representation of the real scene of cognitive world of narrator.

Shorten the Cognitive Distance Between Reader and the Unreliable Narrator

Theory of Mind is the basis for the reader to understand the unreliable narration. Zunshine in her *Why We Read Fiction: Theory of Mind and the Novel* (2006) argued that reading engages and improves mind reading abilities. First, we should acknowledge Nünning (1997) had discovered the cognitive distance between the reader and the text world. Nünning (1997) claimed whether a reader can judge a narrator as a reliable or unreliable is based on his own world knowledge. If a reader did not have the knowledge of mind reading, he cannot understand the reason for the young man in Poe's (2012) novel who always steer the floor and judge the young man as an unreliable narrator. However, the unreliable narration can also help to improve the reader's mind-reading capability. As Turner (2011) claimed that: "Human beings go beyond merely imagining stories or concepts that run counter to the present environment. We can also connect them and blend them to make third mental array" (p. 41). That is, the reader of an unreliable narration can construct the logical plots and reasonable story; this involves the readers' mental blending capability, which is combining the real world and the fictional world of the unreliable narrator to form a relative reasonable world, which was also called the naturalize process by critics. On the contrary, the author (the implied author) deliberately covers the information provided by the narrator, which increases the difficulties of understanding the text. For example, in Joyce Carol Oates's *My Sister, My Love*, the narrator repeatedly retold the scene of the night in which his sister was murdered, the reader had to imagine the scene and guess the possible murder following his revised memory for that. But for his under-age and mental disorder, he presented the reader with fragmented piece of memory and unreadable language of his narration, which seriously hinder the reader to get the reliable information from his fragmented memory. So the readers need put more cognitive effort into the understanding the real meaning of the unreliable narrator and make clear the logical order of the whole story. As Zunshine (2011) proposed that we seem to enjoy being able to exercise this skill: working out the others keep hidden.

To some sense, it is important for how the narrator said, not what he said. For example, Oates' novel *Daddy Love* (2013), the novel was beginning with three different depiction of the scene of kidnapping. Most of readers maybe frustrated by this long and repeated beginning, but when we read the whole novel, we cannot wait to read the second time to find out the reason for the open repetition—this kind of beginning is the imitation of the real people's cognition condition when they faced with dangerous situation. So Oates (2013) depicted the unreliable narrator not in the sense of the content of their telling, but in the sense of the way of their telling, that is, what they say is almost reliable, but how they say is unreliable. In this sense, traditionally, the reader may easily recognize the contradiction in the narration of unreliable narrator. After the cognitive turn of literature study and creation, there are more and more cognitive fictions to challenge the readers' judgment and experience, which directly reveal characters mental experiences, raw sensations, and temporal perspective.

On the other hand, the unreliable narrator provides the reader a new perspective to recognize the world. Most of us see the world from the normal sense, and cannot have the opportunity to see the world from the other's point, especially the abnormal mental world, or the distanced childhood, which may help us to get a good understanding of the work of others' minds. In this sense, the reader is the same to the observer who involves into the feeling of seeing the world from the abnormal perspectives of the unreliable narrator, which is different from their normal life. In one word, the unreliable text provides the reader an opportunity to exercise their mind to experience a different world. "Thus the extended version of our self in the simulated world can experience things, and feel things, that in the everyday world it would not. We project ourselves into the simulated world, and we may come back changed" (Oatley, 2011, p. 23). Reading helps us the change our minds about the world.

In reality each reader, when he is reading, is uniquely reading himself. The writer's work is only a kind of optical instrument which he offers the reader to enable him to discern what without this book he might not have seen in himself. (Proust, 2003, pp. 219-220)

So, as a result of such experiences of mind-reading of unreliable, everyday selves can increase the mental intelligence of others' minds.

So we can get the conclusion for the significance of unreliable narration for the readers: (1) It is resonated with the sympathetic reader's values; and (2) It blurs the reader's source-tracking ability. Most of unreliable narrators may have some cognitive causes for their unreliable narration. For example, although Herbert in *Lolita* (2009) intrigued so many controversial opinions about his behavior, he succeeded in winning the reader's sympathy when he tried to blur the reader's source-tracking ability, in spite of the fact that he may be considered as the rapider of *Lolita*. Literature provides the reader a perfect way to know what the other might think in some circumstance, just as Proust (2003) said: "[the novelist] sets loose in us all possible happinesses and all possible unhappinesses, just a few of which we would spend years of our lives coming to know" (p. 87).

Conclusions

All in all, there is a close relationship between Theory of Mind and the unreliable narrator, that is, Theory of Mind helps us to identify the unreliability in the narration; on the other hand, the unreliable narrator provides now evidences for the study of the Theory of Mind. Palmer (2011) proposed that:

Overall, theory of mind has a projective quality: we cast a model onto another person in ordinary life, or onto a character in fiction. The model can be good, but as we interact more with the person or character, we sometimes find our model to be mistaken in some important respect: an opportunity to improve it, engaging in fiction but usually painful in real life. (as cited in Oatley, 2011, p. 16)

So fiction offers a fuller disclosure than we can obtain elsewhere of the contents of unreliable minds.

References

- Booth, W. C. (1961). *The rhetoric of fiction*. Chicago: University of Chicago Press.
- Faulkner, W. (2009). *The sound and the fury*. London: Vintage Classics.
- Marron, O. (2011). Alternative theory of mind for artificial brains: A logical approach to interpreting alien minds. In P. Leverage (Ed.), *Theory of minds and literature*. West Lafayette: Purdue University Press.
- McEwan, I. (2003). *Atonement: A novel*. New York: Anchor Books.

- Nabokov, V. (2009). *Lolita*. London: The Penguin Books Ltd..
- Nünning, A. (1997). But why will you say that I am mad? On the theory, history, and signals of unreliable narration in British fiction. *Arbeiten zu Anglistik und Amerikanistik*, 22, 83-105.
- O'Connell, S. (1997). *Mindreading: An investigation into how we learn to love and lie*. Oxford: Oxford University Press.
- Oates, J. C. (2008). *My sister, my love*. New York: Harper Perennial.
- Oates, J. C. (2013). *Daddy love*. New York: Atlantic Monthly Press.
- Oatley, K. (2011). Theory of mind and theory of minds in literature. In P. Leverage (Ed.), *Theory of minds and literature*. West Lafayette: Purdue University Press.
- Olson, G. (2003). Reconsidering unreliability: Fallible and untrustworthy narrators. *Narrative*, 11, 93-109.
- Palmer, A. (2011). Social minds in little Dorrit. In P. Leverage (Ed.), *Theory of minds and literature*. West Lafayette: Purdue University Press.
- Phelan, J. (2005). *Living to tell about it*. Ithaca: Cornell University Press.
- Poe, E. A. (2012). *Edgar Allan Poe selected poetry and tales*. M. H. James (Ed.). Peterborough: Broadview Press.
- Proust, M. (2003). *In search of lost time: Finding time again*. (I. Patterson Trans.). London: Penguin.
- Turner, M. (2011). The way we imagine. In P. Leverage (Ed.), *Theory of minds and literature*. West Lafayette: Purdue University Press.
- Wilde, O. (2005). *The picture of Dorian Gray*. San Diego: ICON Group International, Inc..
- Zunshine, L. (2006). *Why we read fiction: Theory of mind and the novel*. Columbus: Ohio State University Press.
- Zunshine, L. (2011). Theory of mind and fiction of embodied transparency. In P. Leverage (Ed.), *Theory of minds and literature*. West Lafayette: Purdue University Press.