SUN MEDIA
BLOEMFONTEIN

A (Tall) Tale of Two Sisters: Integrating rhetorical and cognitive-pragmatic approaches to explore unreliable narration in film

Johanet Kriel

Department of Art History and Image Studies, University of the Free State, PO Box 339, Bloemfontein 9300; email: krielj@ufs.ac.za

There is a sustained debate in the academy about the role of narratology in film studies. This article forms part of this larger debate in exploring the application of the concept of unreliable narration to films, specifically to Jee-woon Kim's little-known but exceptional film *A Tale of Two Sisters* (2003). A dispute surrounding this narratological device has centred on how readers or viewers determine that the narration deviates from diegetic truth. Two major strands of narratology have given divergent answers to this question: the rhetorical approach has been in favour of aligning diegetic truth with an "implied author", while the cognitive approach has called the implied author into question, instead focusing on the viewer's construction of the diegetic truth. This paper investigates the possibility of integrating the two approaches in terms of the viewer's construction of ethical judgements and cued inferences, which would open up a new avenue for considering this narrative device.

1. Background and aim

Unreliable narrators have long been both condemned for their deception and simultaneously adored for the humorous effect and/or "twist" their very deception produces in a novel or film. Just think how different Emily Brontë's *Wuthering Heights* would be without Nelly Dean's colourful narration, while films like Martin Scorsese's *Shutter Island* would hardly have the desired effect if Teddy Daniels were mentally sound. In the former, Nelly is clearly a *narrator* in the conventional sense – we are aware that she is actively telling and thus shaping events for the listener and reader – whereas in the latter, we are rather dealing with a *focaliser* – the viewer is not immediately aware that we are being shown Teddy's specific version of reality. Much has been written about this distinction, but, in terms of this paper, the concepts would be equally relevant whether a conventional narrator is at stake or whether the story is in fact less obviously focalised through a certain character.

Now, while most readers and viewers would be in agreement about the necessity of unreliable narration or focalisation in such beloved novels as *Wuthering Heights* or films like *Shutter Island*, there has been much disagreement in the academy about exactly how readers and viewers detect this device and resolve the resulting inconsistencies within the narrative. Per Krogh Hansen (2008: 2), a narratologist specialising in film studies, summarises the recent trends as follows:

> As is probably well-known, there have been two major concerns here [in the application of the unreliable narrator to film] recently: the one being that of rhetorical criticism with a focus on the ethical positioning of the reader. James Phelan is a leading figure in this connection (Phelan 2005). The other is that of cognitive studies with an interest in the reader's constructive involvement in determining the unreliability which (at least until 2005 (Nünning 2005)) have had Ansgar Nünning as the prime mover.[1]

These two positions have often been placed in contrast to one another, because the rhetorical approach focuses on the reader's ethical positioning in relation to a kind of speaker or "implied author" who is representative of the actual author and his/her values, whereas the cognitive approach focuses on the reader's construction of meaning as cued by the text. Nevertheless, it seems to me that the two sides of the debate need not be mutually exclusive. After all,

---

1   Indeed, this would still be the case, as is evidenced by the recent publication of Nünning's most recent book – *Unreliable Narration: Studien zur Theorie und Praxis unglaubwürdigen Erzählens in der englischsprachigen Erzählliteratur* (2013).

both consider narration to a larger or lesser degree to be a communicative act and certainly, the reader's *ethical positioning* influences his/her *construction* of the narrator as unreliable, and vice versa. Thus, although I mostly follow the cognitive(-pragmatic) approach due to its versatility, I incorporate aspects of the rhetorical approach to understand the viewer's construction of diegetic truth and detection of unreliable narration.

My aim is thus to delineate a relevant and novel definition of unreliable narration that incorporates aspects of the conflicting perspectives. To do this, I firstly look at how one could apply the notion of a narrator, which originates in literary studies, to the cinematic medium. Then, I consider the origins of the term "unreliable narration" in the writings of Wayne C Booth, one of the earliest and most influential scholars in narratology. I consider the advantages and disadvantages that his and his followers' rhetorical approach presents for the purposes of understanding readers'/viewers' construction of the narrator's reliability. Specifically, I look at the much-debated concept of the "implied author", and at the ethical responses of the reader/viewer to the unreliable narrator/focaliser. I then turn to the cognitive and pragmatic approaches. These have often positioned themselves in resistance to the rhetorical approach in favour of a more constructivist outlook, which has its roots in cognitive science and communication studies. Taking my cue from Warren Buckland (1995: 55), I thus similarly investigate "the question of intelligibility in the cinema – i.e. how does the spectator make sense out of the basic plot structure of Hollywood narrative film" and how this influences the debate surrounding unreliable narration in films. Accordingly, while I mostly build on Buckland and Daniel Barratt's cognitivist approaches, I cross-pollinate their theories with Booth and others' rhetorical approach. The cognitivist approaches are pragmatic and the least problematic to apply, because their arguments tend to centre on the more easily demonstrable, lower-level cognitive processes, but the drawback is that they thereby tend to exclude those higher-level mental and emotional responses which are harder to gauge, such as assessing the moral reliability of the characters. In terms of the latter shortcoming, I argue that one could integrate aspects from the rhetorical approach to better understand the viewer's estimation of a narrator's moral reliability and thus his/her veridicality without turning to the problematic notion of the implied author.

In order to ground my argument, which otherwise runs the risk of becoming too abstract and theoretical, Jee-woon Kim 2003's psychological horror *A Tale of Two Sisters²* is analysed throughout the paper. I chose this film because of its adept inclusion of unreliable narration and because it is an outstanding film in terms of its plot and cinematography about which little has been written from an

---

2   Hereafter abbreviated to "*AToTS*".

academic point of view. Given the relative obscurity of the film in comparison to other films featuring complex narration, I begin with a brief synopsis.

2. Synopsis of *A Tale of Two Sisters*

Befitting the psychological horror that it is, the film opens with a psychologist/psychiatrist waiting for and then engaging with his young, female patient named Su-mi, who is semi-catatonic and guided into a chair by a nurse. He asks her questions about her past and family that suggest that she has suffered some lapse of memory. Along with the setting, this should already be a clue that unreliable focalisation may be at stake. Su-mi is unresponsive until he shows her a picture of her family, at which she slowly lifts her head and then stares out of the window.

The film then cuts to a scene of a car driving through the countryside and we see Su-mi arriving at home with her father, Moo-hyeon, and Su-yeon, her timid, sweet-natured, younger sister. This is the first ambiguity in the film, because this scene could suggest either a flashback/analepsis (i.e. Su-mi is thinking back to an earlier homecoming prior to her hospitalisation) or a flash forward (i.e. some time has passed and we see Su-mi returning home from the hospital). Both deductions prove to be equally valid, but for simplicity's sake, I will regard this as a flashback.

After playing outside and sitting together on a nearby wharf, the girls enter the dark and eerie house where their stepmother, Eun-joo, startles them. She feigns a friendly, welcoming attitude, yet the girls seem frightened of her and the interaction is unsettling, especially the pointed way in which she remarks on Su-mi's mental health. As the film progresses, the atmosphere becomes stranger and the relationship between the girls and the stepmother worsens to the point of physical violence. The house is apparently haunted by the staple of many Asian horror films – a pale woman with long dark hair hanging over her face. A friend of the family also sees the apparition, so that the viewer assumes that the ghost is diegetically real. In the same way, we assume that Su-yeon's mysterious bruises are real and that Eun-joo is the source thereof, since she is portrayed as increasingly manic and brutal and we realise that she is on (presumably psychiatric) medication. Eun-joo finds her murdered pet birds in Su-yeon's bed and it is initially unclear whether the ghost is responsible or whether the girls killed them as an act of revenge. Eun-joo punishes Su-yeon by locking her in a wardrobe that she is clearly terrified of. Su-Mi comes to her sister's rescue and confronts her father, who for most of the narrative has remained aloof and apparently oblivious to what has been happening between the other three family members.

This is when the unreliable narration is suddenly revealed: the father begs Su-mi to stop "making trouble" because Su-yeon is dead. At the same time, the

younger sister is clearly visible to the viewer and Su-mi; in fact, Su-mi weepingly tells Su-yeon that this is not true while the latter begins to scream. This unsettling scene suddenly cuts to another in which Su-mi finds a bloody bag containing what she believes to be Su-yeon's body. A physical struggle ensues between Su-mi and Eun-joo, which ends with the woman knocking the girl unconscious. When Su-mi regains consciousness, we learn through an analeptic montage that the "Su-yeon" and "Eun-joo" we have seen up to this point have merely been figments of Su-mi's imagination. It thus becomes clear that our view of the events has been unreliably focalised through Su-mi. Su-mi is psychologically disturbed and has in fact been assuming the personas of Su-yeon and Eun-joo, while the real Su-yeon is dead and the real Eun-joo has been living somewhere else.

Thereafter follows a flashback scene to an even earlier time which is interspersed with scenes from the film's present time. It is revealed that Moo-hyeon was apparently having an affair with Eun-joo, who was then the live-in nurse taking care of his mentally ill wife. One afternoon, Eun-joo has a brief altercation with the two girls and Su-yeon flees upstairs to seek comfort from her mother. Su-yeon falls asleep in her mother's lap, but when she wakes up, she finds that her mother has hung herself in the same terrifying wardrobe shown at an earlier point in the film. Su-yeon frantically tugs at her mother's body, upsetting the wardrobe in the process, so that it and her mother's body fall and trap her beneath.

We then return to the film's present, and see the real Eun-joo visiting Su-mi in a psychiatric hospital. As in the earlier imagined homecoming scene, Eun-joo is coyly yet eerily friendly and Su-mi responds by painfully grabbing Eun-joo's arm in a tight grip, upon which Eun-joo's mean streak flares up. Eun-joo is further cast as morally suspect since we see Moo-hyeon and the doctor apparently discussing her. Eun-joo and Moo-hyeon drive back to the house in silence. Once there, the female phantom, who is probably the ghost of the girls' mother, attacks and presumably kills Eun-joo. There is a quick cut from Su-mi looking up from her hospital bed to Eun-joo screaming in the house.

We are then shown the rest of the flashback. Eun-joo, who has heard the thud of the falling wardrobe, races upstairs and sees Su-yeon's arm flailing under the wardrobe. Her first instinct is to leave the room, but presumably her conscience gets the better of her and she turns back. Unfortunately, she encounters a still irate Su-mi who rudely informs Eun-joo that she does not belong in this part of the house. Eun-joo snaps at Su-mi that she will regret this moment and Su-mi retorts that she is already regretting it and then storms out of the house. It becomes clear that because of this Eun-joo did not turn back to help Su-yeon, and we realise that the little girl has suffocated to death in the meantime. The

film ends with a shot reminiscent of François Truffaut's *The 400 Blows*: a black and white freeze frame of Su-mi's face as she stomps away from the house. The credits of the film roll over a shot of Su-mi sitting by the wharf on her own; an eerie inversion of the earlier imagined scene in which she and her sister dangled their legs in the water.

While any synopsis naturally falls short of the actual film, I hope to have conveyed how central unreliable narration is to this particular film. Moreover, because we are finally uncertain as to what actually did occur in terms of diegetic truth, we cannot be truly certain which parts of the film were unreliably focalised. For example, we are left to wonder whether it is the "real" Eun-joo we see in the house at the end of the film, or simply a psychic projection by Su-mi; whether Eun-joo actually abused the girls; and whether the diegetic present of the film has now gone full circle to the opening scene or whether this is a different visit to the hospital. The film's unreliable narration thus makes for an interesting case study, particularly in terms of how Eun-joo is initially depicted as the typical "evil stepmother". While there is much more one could write about the film itself (e.g. in terms of psychoanalysis), this is not possible here. Instead, I restrict my focus to the perception of narrative levels and unreliable narration. I thus move on to my theoretical framework in terms of unreliable narration.

3. Narrators in film

First of all, one has to admit that narration functions very differently in films and other moving images than in literary narration, mostly due to the lack of a clearly definable narrative voice. Several narratologists like Volker Ferenz (2005: 135) thus argue that it is necessary to narrow the theoretical scope to include only films which feature a "pseudo-diegetic character-narrator" who "appears to be in the driving seat of the narration" under the category of unreliable narration. However, this reduces our investigative scope to exclude not only films by the likes of David Lynch and David Cronenberg but also literature by writers such as J M Coetzee and Michel Houellebecq, in which there is no clear narrator or focaliser yet where unreliable narration is still at stake. Hansen (2008: 7) struggles with this same problem and reaches a conclusion which does not entirely resolve the issue, but which is sufficient for my purposes here:

> One might claim that cases like these are more adequately described [...] as "unreliable focalization", but this of course depends on how one understands third-person narration: whether it is characterized by "someone" (the extradiegetic narrator) telling *about* the incidents and characters, or whether it is the third person *per se* who narrates through the means of

> free indirect discourse, covert narration, etc. Many third-person narrations have restricted the focalization to one character, and insofar as the extradiegetic narrator does not show any superior knowledge (signs of omniscience) or delegates the perspective to other characters, it makes good sense to consider this mode as "covert first-person narration" or "third person proper" insofar as the extradiegetic narrator does not have any or only a little significant function. In cases like these, we might very well observe unreliable reporting or judgement of the narrated events. [original emphasis]

In terms of unreliable narration in films, I suggest that it could prove fruitful to broaden this definition of unreliable narration or focalisation in films to include instances where the film may in fact feature both kinds of third-person narrators as in *AToTS*. We encounter both "covert first-person narration" (limited to Su-mi's perspective), and the more conventional third-person narrator *per se* (neither Su- mi nor the figments of her imagination are present, specifically, in those scenes which feature the "real" Eun-joo but not Su-mi). Interestingly, because the first and last shots suggest that Su-mi is the focaliser, one could argue that the reliability of the entire narrative (including both the framing and the embedded narrative) is compromised. However, since this would make the film impossible to analyse in terms of diegetic truth, I opt for the more conservative interpretation; namely that only part of the film has been unreliably focalised and that one can unravel and rearrange the narration to come to some kind of diegetic truth. The issue of diegetic truth takes me to the rhetorical approach's central concept of the implied author.

4. The rhetorical approach: the question of the implied author and moral norms

Wayne C Booth first coined the term "unreliable narrator" in his seminal book *The Rhetoric of Fiction*, in which he stated that a narrator is "*reliable* when he[/ she] speaks for or acts in accordance with the norms of the work (which is to say, the implied author's norms), *unreliable* when he[/she] does not" (1983: 158- 9, original emphasis). With regard to the "norms of the work" and those of the implied author, Booth maintains that we must look at the "*moral* and *intellectual* qualities of the narrator" (1983: 158, emphasis added) in order to determine whether or not he/she qualifies as "fallible or unreliable" (terms that Booth, in fact, employs synonymously most of the time). Clearly, the further the narrator strays from the "norms of the work", the more unreliable he/she becomes. There are thus two concepts underlying Booth's original definition which have been widely contested, namely, the implied author and the ethics of fiction.

One of the most vocal detractors is Ansgar Nünning (1997, 1999, 2008 & 2013), who criticises the concept of the "implied author" and the "secret communication" that takes place between him/her and the reader(/viewer) as being built on uncertain theoretical ground. To some extent, I agree with Nünning (1997: 87) that "the implied author's norms are [for all practical intents and purposes] impossible to establish and that the concept of the implied author is eminently dispensable". He also points out the weak points of several other critics in this regard, including Chatman, Prince, Rimmon-Kenan, Riggan, Wall and Yacobi – and I would add Currie (1995), Phelan (2007) and Lothe (2000) to this list – all of whom rely implicitly or explicitly on the implied author[3]. However, a problem arises in that no wholly satisfactory account has been provided in terms of where the implied author comes from. For one thing, how does a reader (re)construct an implied author? If it is from textual signals alone, why do we feel the need at all to construct an entity which exists somewhere between the text and the actual author? If it is not from textual signals alone, where does one draw the lines between the narrator, the implied author and the actual author? And how does this account for the differing interpretations of readers, something which also troubled Booth in his writing on the topic? I would also add that the implied author brings us to another contentious issue: authorial intentionality; i.e., is our implied author one and the same implied author as the actual writer or filmmaker *intended*?

The terrain becomes even more treacherous in films, where the viewers have to employ very different kinds of schemata to make narrative sense of what they see and hear. Booth himself admits to this difficulty in a 2002 article on *American Beauty* (Sam Mendes, 1999). After giving stimulating insights into changes made to the script during production and probing various other facets of the film, Booth (2002: 129) concludes: "For me, there *is* a center to *American Beauty*, a center that can never be adequately formulated in words: it is found in the creative energy that hundreds of people put into its production." He then goes on to basically describe the "message" or "meaning" of the film (what David Bordwell (1991) would call the "implicit" or even "symptomatic" meanings). Although Booth's argument is stimulating and refreshingly self-reflexive, it makes for rather slippery theorising if we have to locate the implied author in a conflation between the entire production team, the original script and the meanings we interpreted for the film. Thus, while I am not quite as radical as Nünning in my rejection of the notion of the implied author (especially where the author's intentions are relevant and can be more certainly established), in general, I find it an unnecessary complication in the estimation of unreliable narration.

---

3   For a more detailed overview, see Nünning (2008).

Without this authority to contradict, however, how do we recognise an unreliable narrator? I would agree with Nünning (1997: 87) that unreliable narration can be found at the crossroads between the narrator's *view* or rather *focalisation of diegetic truth* and the *reader/viewer's*[4] *understanding of diegetic truth*. In addition, it is not enough to view a film in terms of being a "totality of textual signals", to borrow Greta Olson's (2003: 99) phrase. Nünning (2008: 45) thus asserts:

> What is needed instead is a pragmatic and cognitive framework that takes into consideration the world-model or conceptual information previously existing in the mind of the reader or critic. It is necessary to take into consideration both the world-model and norms in the mind of the reader and the interplay between textual and extratextual information.

The viewer's cognitive frames of reality are thus considered alongside the text. This shift of focus also allows space for the variable responses of viewers without having to rely on the implied author.

Furthermore, Nünning (1997: 89) for one criticises Booth's ambiguous use of the word "norms" as implying *both* epistemological and moral standards. He suggests that "unreliable" should relate to the former (cases where the narrator incompletely or incorrectly renders narrative aspects) and "untrustworthy" should relate to the latter (the narrator makes morally unsound judgements of the narrative events)[5]. While these distinctions prove expedient on paper, they are not always distinct, instead forming a kind of continuum between unreliable and untrustworthy. This is due to a correlation between characterisation and unreliable narration, because "[t]he general effect of [...] unreliable narration consists of redirecting the reader's attention from the level of the story to the speaker and of foregrounding peculiarities of the narrator's psychology" (Nünning 1997: 88)[6]. In *AToTS* this occurs when our attention is redirected from the tale of two sisters and their stepmother to the characterisation of Su-mi as psychologically disturbed and hence as potentially both unreliable and untrustworthy.

One could thus level many criticisms at Booth's work (indeed, Booth himself does so in the afterword to the second edition, 1983: 422-3), but he does provide relevant insights as well. Citing Booth, I would add that the narrator's moral

---

4   For a more detailed overview, see Nünning (1999:58).

5   Olson (2003) adds "fallible" as another useful designation, indicating cases where a narrator has limited *access* to the diegetic truth. However, I will not investigate this further here, since it relates more to diegetic events (e.g. a narrator being absent from a specific event) than to our construal of the narrator as unreliable *per se*.

6   For more on this notion, see Bernaerts (2008:92-96).

values, as resulting from but also as distinct from his/her psyche, do for the better or worse contribute to our judgement of a narrator as unreliable. Moreover, the viewer's deeming the narrator as unreliable often goes in hand in hand with a *moral* value judgement, and vice versa – once a narrator has proven to be *ethically* and/or *psychologically* unsound (i.e. untrustworthy), we will be all the more ready to expect his/her narration to be *unreliable* (e.g. the narrator of *Fight Club* and Humbert Humbert in *Lolita*[7]). While Nünning admits to this possibility, he is against incorporating it in our theorising as it is too difficult to determine. I would however maintain that our moral engagement with the narrator/focaliser is too important to leave out. As Hansen (2008: 4) succinctly puts it:

> Even though Ansgar Nünning is right in claiming that we in general are not aware of our own presuppositions and moral convictions, this does not mean they have to stay unacknowledged. A part of our task as professional readers is to reflect on our readings and interpretations – i.e. to consider the framework within which they are established and perhaps reframe them – and to understand the ethical issues involved, which of course suggests an obvious bridge to the rhetorical approach to unreliable narration.

AToTS is an especially interesting case as it is different from cases such as *Fight Club* and *Lolita.* While we accept that Su-mi is psychologically unsound and thus an unreliable narrator, this does not necessarily make her morally untrustworthy. Instead, we are prompted to view Eun-joo as the "evil stepmother" archetype, and the two sisters as the "innocent victims". As such, we are in fact initially inclined to trust Su-mi as a focaliser and to view the strange events in the house as due to either Eun-joo's strange behaviour or to the ghost's supernatural whims. As a result, we miss subtle clues that unreliable narration may be at hand; e.g. the fact that their father never addresses Su-yeon directly. In this case, the ethical response of the viewer is vital if the "trick" of unreliable narration is to work on us. To better understand how the film's unreliable narration works, I turn to the cognitive and pragmatic approaches.

---

7   James Phelan (2007: 225) makes an interesting case that in certain instances, "although the authorial audience recognizes the narrator's unreliability, that unreliability includes some communication that the implied author – and thus the authorial audience – endorses". While I agree, I do not investigate this here, since it is not applicable to AToTS.

5.   Cognitive-pragmatism: intratextual signals and extratextual inferences

In order to better understand unreliable narration, one has to understand how the reader/viewer constructs diegetic truth. Nünning (1997: 95-6) claims that the factors for determining the unreliability of a narrator are located both intratextually (in the "totality of textual signals") as well as extratextually (in the reader's comprehension of these collective signals). He argues that the reader applies four referential frameworks in comparing what is located inside the film with what he/she has learnt outside of it. The first three relate to the reader's world-knowledge; namely, empirical verisimilitude, cultural norms of psychological normality, and ethical and moral standards; while the fourth entails those frames of reference which are specific to constructed narratives, e.g. genres and character archetypes (Nünning 1997: 100-1). These referential frameworks which we use to make sense of narratives are similar to those that we employ to make sense of our own lives and as in real life, we accept that for the most part the default mode of a text is "reliable". When either our sense of reality or the reliability of a narrative is challenged, we apply our referential frameworks to try to find a reason (either psychological, moral, or even artistic in the latter case) for this contradiction. Warren Buckland and Daniel Barratt both offer useful insights into how a reader/viewer resolves these contradictions in terms of unreliable narration.

5.1   Warren Buckland on feasible and defeasible inferences

Buckland follows a pragmatic, communicative and cognitive approach[8], because, he claims, "the spectator's understanding of a narrative film depends upon his/her information processing activity, which consists primarily of the generation of non-demonstrative inferences – that is, inferences primarily deducible from contextual information, rather than from the semantic content of the film" (1995: 55). This assertion is based on the pragmatic model of communication, which is based on the observation that actual communication is neither complete nor perfect. Similarly, the communication system that is employed in cinema is as devoid of a "failsafe semantic algorithm" as any other form of human communication. "The message the sender wishes to send cannot be automatically encoded into an utterance and then automatically decoded by the addressee", which is not to say we should abandon semantics and code models, but this "severely limits their descriptive adequacy" (1995: 56) in the case of films. Buckland's pragmatic approach thus helps us to understand how gaps open up in

---

8   In this respect, Buckland's approach is largely adapted from Sperber and Wilson (1986).

these communications that allow the possibility of unreliable narration slipping in without initial detection by the viewer. Furthermore, through analysing how we close these gaps, Buckland addresses an issue which Nünning does not; namely, the role of viewers' expectations in drawing certain inferences from films. Buckland argues that our expectations rely on the inferences we form and vice versa. In *AToTS*, one of the major expectations which shapes our inferences is that we are about to view a psychological horror film. We thus might expect to see a morally reprehensible character who is driven by pathological urges and might expect some supernatural phenomena.

Buckland also contends that our inferences can be either demonstrative or non-demonstrative. The difference between the two lies in the fact that demonstrative inferences are "deducible solely from the semantic content of an utterance" and are easily "proved on the basis of a set of deductive rules" (1995: 56), for example Eun-joo saying "Welcome home". Non-demonstrative inferences, on the other hand, cannot be construed from the semantic content of an utterance only and cannot be "proven" according to a set of rules. Instead, they are spontaneously drawn from contextual information needed to "complete" communication (insofar as that is possible from a pragmatic standpoint). In viewing films, we thus usually rely on non-demonstrative inferences. For example, we would form the non-demonstrative inference that not only is Eun-joo ill but that she is ashamed of her illness because she looks embarrassed when Moo-hyeon gives her a pill to take and neither of them says anything.

Another useful aspect is that non-demonstrative inferences are deemed as being "successful" or "efficient", rather than logically "valid" or "proven". This assessment is based on two stages: hypothesis formation and confirmation/disconfirmation (1995: 56-7). Thus, what makes non-demonstrative inferences particularly useful here is that in unreliably narrated films, certain intratextual signals as well as intertextual signals (such as generic framing and allusions to other texts) prompt us to form certain hypotheses and "complete" them according to our frameworks, only to have them disconfirmed later during our viewing. This two-fold process is vital if the "trick" of the film is to work on viewers and if we are to perceive narration or focalisation as unreliable at all. Indeed, in *AToTS* the first and foremost non-demonstrative inference that we are prompted to make is that the flashback which we witness for most of the film is diegetically truthful; that it is an analepsis which is reliably and extradiegetically narrated by Su-mi who is then in a psychiatric institute. This is verbally signalled by the psychiatrist's words, "Tell me about what happened that day – you should be able to remember it clearly," and Su-mi's staring out of the window as if she is indeed remembering something. The shot fades into a point-of-view (POV)-shot of a summer landscape passing by a car window with beautiful extradiegetic music

playing, indicating that we are witnessing Su-mi's memories of a better time. We thus disregard the earlier signals that she might be mentally unstable and thus an unreliable narrator; instead, we trust her recollection of events. The question is why we are not consciously aware of this distortion even though the framing narrative provides us with some early clues, as is the case with *The Sixth Sense*, as discussed by Hansen (2008), and *A Beautiful Mind*, as discussed by Barratt (2009). Because film is a hypotypic (a highly mimetic) medium (cf. Hansen 2008: 10) and our default viewing mode is that what we are seeing is diegetically real, we assume that Su-mi's flashback is reliably narrated (which entails that the characters and events are real). This assumption is promptly reinforced when we see Su-mi and Su-yeon from a vantage point somewhere inside the house. This technical detail signals that the analepsis is not restricted to Su-mi's viewpoint but that a more omniscient and therefore presumably objective and reliable kind of focalisation is at stake. This is further supported by the representation of certain diegetic events which Su-mi could not have witnessed herself (such as her father's phone conversation or the interactions between him and Eun-joo).

Another important factor which deflects our attention from the possibility of unreliable narration is the stepmother and her relationship with the girls. Firstly, we infer that Eun-joo is immoral and psychologically unsound and thus unreliable, and secondly that by extension the girls are innocent victims and thus reliable. These inferences are reinforced by the intertextual framing of the film. The original title of the film is 장화, 홍련, literally translated as "Rose Flower, Red Lotus" (Wikipedia 2014⁹). It refers to a well-known Korean folktale in which a cruel stepmother executes an evil scheme against her two stepdaughters, which includes putting a skinned rat in the bed of one of the girls in order to make it look as if she had a miscarriage, so as to foil her wedding plans. Her scheming leads to the death of both girls and their ghosts haunt the local area after that. The film includes several implicit references to the tale, such as the small piece of meat Su-mi discovers in the refrigerator which may be a reference to the skinned rat and the bloodstain on the bed from Su-yeon's first menstruation. The main reason I mention the folktale is because if we were familiar with the folktale or the other films derived from it, we would already have certain expectations and we would be more likely to form certain inferences. This framing also reinforces two additional inferences which make it hard for us to spot the unreliable narration; namely that all the characters are diegetically real and that the house is haunted by an actual ghost. The viewer is thus shown a "real" ghost, which conforms to our generic expectations and which contrasts starkly with a character like Su- yeon, who is an adorable young girl and we thus infer is not a ghost. The

---

9   Unfortunately the only English-language source available on the original folktale.

house and the stepmother are also cast in eerie contrast to the two girls. When Su-mi first looks at the house, the music stops abruptly and the soundtrack remains silent during several intercuts with low-angle shots of parts of the house shown to be empty and in shadow. The eerie sensation created by these shots and the interaction with Eun-joo would enforce our generic framing that we are seeing a kind of haunted house with a deranged parent as in films like *The Shining* and *The Amityville Horror*.

To understand how our inferences can later be disproven without refuting the film's logic, Buckland's explanation of "relevance" and "defeasibility" is expedient. To put it very basically, the addressee of any communication decides whether and to what degree new information is "relevant" (makes sense in terms of the situation). This is based on whether the information has a "contextual effect" – if it is "is new and relates to information already acquired by the addressee" – and whether the "processing effort" required to synthesise this information is relatively small (Buckland 1995: 57). In other words, the inferences we usually generate are those that "[produce] the optimal contextual effect with only a small (adequate) amount of processing effort" (1995: 57); and this appears to be based on evolutionary efficiency. To return to my example of *AToTS*, for the (first-time) viewer to process each piece of new information in the embedded, analeptic narrative as a possible distortion would amount to a large processing effort in comparison to the contextual effect. Because we would constantly have to question all that we see, it is easier to believe that the stepmother is cruel, that the ghost is real, etc. and to deem these inferences as "relevant". Barratt (2009: 78) would add that thematic and character structures are also influential on what we deem to be relevant. In *AToTS* we are prompted to cast the stepmother and ghost as the "bad guys" and the girls as the "good guys", so that we miss minor clues indicating unreliable focalisation.

Buckland's notion of "defeasibility" helps us to understand how a film could initially compel us to make one set of inferences, only to later disprove them. Basically, defeasibility is the principle that inferences "can be cancelled without creating a logical contradiction" (1995: 58) within a narrative. To explain defeasibility in everyday communication, Sperber and Wilson (1996: 260) provide the example of an addressee replying, "Coffee would keep me awake" when asked, "Will you have some coffee?" Based on our contextual assumptions, we could relevantly infer either that he/she wants coffee (he/she wants to stay awake) or not (he/she does not want to stay awake), and neither would create a logical contradiction within the utterance itself. However, when we realise in the light of new contextual assumptions (e.g. the person wants to stay awake to study), that our earlier inferences were "defeasible", we are able to create new inferences without creating a logical contradiction within the utterance. Unreliably narrated films often rely on this distinction; they rely on the viewer creating a relevant

inference based on the cues in the film, and then prove our inferences defeasible later on. To better understand why we create these defeasible inferences, we need to look at how viewers form contextual assumptions.

Buckland (1995: 59-60) refers to neo-formalist David Bordwell to remind us that by constructing the *fabula* of *any* film, we make certain inferences based on our contextual assumptions which in turn arise from the everyday cognitive and genre-specific schemata we apply while watching the film. In terms of relevance, Buckland (1995: 61) thus concludes that the *fabula* of most films are usually clearly and reliably narrated in terms of cause-and-effect relationships. We accept the minor "complications" and gaps generated by the *syuzhet* because they keep our interest piqued until the end of the film. Buckland (1995: 61) thus endeavours to build on Bordwell's constructivist "emphasis on the procedural, top-down, defeasible nature of the perceiver's activity in information processing" through continually applying schemata and the relevance principle, to better comprehend how we create, revise and abandon inferences. In the case of my paper, *certain* unreliably narrated films contain signals that in conjunction with the viewer's schemata, prompt him/her to create relevant inferences which are for some time sustained but which are ultimately defeasible, making the narrator unreliable without creating a logical contradiction in the film itself. Accordingly, we are able to sustain for some time – I would say until the climactic scene when we learn that Su-yeon has been dead all along – the inference that the analeptic narrative is diegetically true. When we do eventually learn that this inference is defeasible since Su-mi has been an unreliable focaliser, no logical contradiction is created in the film itself which has merely shown us Su-mi's *account* of diegetic truth. Daniel Barratt has paved the way to understand the role of defeasibility in unreliably focalised films even better.

5.2   Daniel Barratt on the deficiency of attention and memory

To further expand our understanding of a viewer's construction of unreliable narration in film, Barratt (2009: 63) provides two further psychological facets in film viewing that have a decided influence on the non-demonstrative inferences we form and maintain. The first is attention, which is understood as both limited and serial in nature, and the second is memory, which is understood as being more impressionable than we would like to admit.

Firstly, to understand the limitations of our attention, Barratt explains the "primacy effect" and the "priming procedure" along with the effect of new information and emotion on the viewer's perception. The "primacy effect" entails that we assign more importance to what we encounter or infer first (first impressions *matter*), while the "priming procedure" allows this effect

to persist throughout our viewing (first impressions *last*). Thus, for better or worse, our *initial* non-demonstrative inferences shape how we view the rest of the film. For instance, once we believe that the stepmother and Su-yeon are diegtically real, we will sustain that belief until it is directly challenged.

In addition to this, Barratt explains that our attention can be affected by informational and emotional "loads". The informational load basically refers to the fact that at certain stages of the film's narrative, especially the beginning, we are too preoccupied with constructing the basics of the *fabula* to notice details which might make our inferences defeasible. Thus, we are too busy trying to understand what has put Su-mi in the hospital to consider that the flashback might be unreliably focalised. Emotional loads could refer to any kind of emotion which is evoked and which could deflect our attention; for example, at the beginning of the flashback, the pleasant emotion that is evoked by the beautiful scenery and music and the playful innocence of the two sisters strengthens the inference that they are the "good guys". One of the most powerful emotional loads however is fear. In psychological horror films such as *The Sixth Sense* and *AToTS*, fear is especially effective in distracting us from early indications that one of our inferences may be defeasible. This is because fear creates a kind of tunnel vision around the threat and its effects "tend to persist, decaying slowly over time" (2009: 67). Barratt thus (to some extent) addresses a concern which Buckland (1995: 65-6) mentions; namely, "the question of the relationship between logical comprehension and emotional response" and whether they compete with or complement one another. In fact, this is part and parcel of a larger complaint which has often been levelled against cognitivism and linguistics; that they reduce human beings to "purely rational communicators, interested only in the efficiency and productivity of information processing" (Buckland 1995: 65-6). Barratt partly resolves this criticism by indirectly answering Buckland's question: the answer is *both* – our emotional response can complement *and* compete with our logical comprehension of a scene. In psychological horrors like *AToTS*, fear can be triggered to complement as well as interfere with our comprehension of a scene. For example, the fear triggered by the visions of the ghost helps us to comprehend the fear that the two sisters have of the house. However, it can also misdirect our attention. For example, the vision of the blood (and hand!) running down the ghost's thigh as she looms threateningly over Su-mi leads us to infer that the coincidence of Su-mi's, Su-yeon's and Eun-joo's menstrual cycles is caused by supernatural forces, instead of allowing us to recognise a clue that they might in fact be the same person.

In addition to the influence of our limited and easily misdirected attention, our memory of a certain scene is also tainted by the assumptions we held at the time of viewing it. Barratt (2009: 67-8) illustrates this with another instance

from *The Sixth Sense*: if we assume that Dr Crowe is alive when he commences his sessions with Cole, we tend to remember that Dr Crowe and Cole's mother interacted at Cole's house, which in fact they did not. Basically, viewers are usually more "discerning" at the beginning of a film/sequence than at the end and moreover "human cognition tends to be conservative", which means that "[i]n the absence of [relevant] contradictory information, we usually [...] go for the most obvious interpretation of events" (2009: 68). In other words, we trust our initial impressions of a reliable focaliser in *AToTS* and tend to recall scenes incorrectly, for example, we remember the homecoming scene as if Hoo-yeon addressed both girls. In fact, according to Babylon's online dictionary (2011), 너, 네가 and 당신 both mean "you" in the singular *and* plural second person (as in English). Thus, whenever the father addresses Su-mi as "you", we could have assumed that he was addressing her *and* Su-yeon/Eun-joo, as we might have done in any English-language film.

Finally then, what we find is that films like *AToTS* feature several clues to unreliable focalisation but that they apparently rely on our incomplete and conservative attention and faulty memory to create relevant but ultimately defeasible inferences. If this were not the case, the "trick" of *AToTS* and comparable films would not work on us the first time we view them. When we do become aware of how the film has "tricked" us, however, we may well be tempted to watch the film again; only to realise that there is no logical contradiction in the film itself but it is due to our own perception, which was carefully misguided by the film, that we could be "tricked".

6. Reflection and conclusion

In reflection, when analysing films which feature unreliable narration or unreliable focalisation, it has been especially useful to lean on the cognitive approach in order to explore how the viewer creates meaning through intratextual signals, intertextual and generic framing, and extratextual cognition. As such, a cognitive-pragmatic approach can potentially provide a more sound method for analysing the viewer's construction of diegetic truth than the rhetorical approach can. Buckland's delineation of non-demonstrative inferences and (de)feasibility have been particularly constructive to elucidate how films can compel us to draw specific erroneous conclusions, e.g. that certain characters are diegetically real. In films with a "twist" ending, like *AToTS*, one of the inferences which viewers are initially (mis)guided to form is that the narration or focalisation is indeed reliable. In addition, films like *AToTS* and other such "puzzle" films cleverly exploit our fallible memories and easily distracted and serialised attention span in order to produce this "twist" effect without creating a contradiction in the narrative itself.

Thus, when we read the book or watch the film again, we can have the strange pleasure of realising how we have been duped. I have found that psychological horrors are particularly well-suited for unreliable narration, because supernatural phenomena, fearful scenes and psychologically unsound characters are part and parcel of the genre.

While the cognitive-pragmatic approach thus provides a strong analytical tool, it can still be improved if we also turn to the rhetorical approach. In terms of the latter, the rhetorical approach offers the possibility of exploring viewers' higher emotional and intellectual reactions to narratives and characters, which the cognitive approach tends to shy away from. If a filmmaker were thus to include plot and characterisation cues which compel us to condemn a certain character or action, while compelling us to sympathise with and trust another character, we would be all the more easily misled to initially trust an unreliably narrated or focalised reality, thus ensuring that the "twist" would really work on us. In *AToTS* it is especially useful to understand why we are more compelled to view young Su-mi as an innocent victim and thus to deem her focalisation as reliable, in juxtaposition to her morally repugnant stepmother, who, despite her moral flaws, is evidently a reliable focaliser. Integrating the two approaches thus allows us the chance to more fully understand the working of unreliable narration in puzzle films.

In terms of possible future research, I would propose that one could continue this line of invesitgation to understand how unreliable narration might challenge the very construction of diegetic truth by the reader or viewer. For, once the unreliable focalisation and the deficient and susceptible nature of our attention and memory is revealed, we may feel compelled to question the relationship of the narrated levels to the diegetic truth as a whole (i.e. how can we be sure that only this one part of the story is unreliably narrated?) and we might even reconsider our own perception of reality beyond the text (i.e., if I was so easily duped by a film, what else is there that I am missing?). In *AToTS*, the unreliable narration has a far-reaching effect and herein lies the crux and the lasting effect of the film: we can never be certain what *exactly* constitutes its diegetic truth (e.g. is the final scene of the real Eun-joo and the ghost diegetically true or not?), so that in the end we can associate with Su-mi's uncertain relationship with reality more empathically than we might have liked to. Alas, I must leave this line of enquiry for future research.

Bibliography

Babylon Korean to English translator (2011) http://translation.babylon.com/
korean/to-english/ [view date: 10 October 2014]

Barratt D (2009) "Twist blindness": the role of primacy, priming, schemas,
and reconstructive memory in a first-time viewing of *The Sixth Sense*. In:
Buckland W (ed.) *Puzzle Films: Complex Storytelling in Contemporary Cinema*.
Chichester: Blackwell-Wiley.

Bernaerts L (2008) "Un fou raisonnant et imaginant": madness, unreliability and
The Man Who Had His Hair Cut Short. In: D'Hoker E & Martens G (eds.) *Narrative
unreliability in the twentieth-century first-person Novel*. Berlin: De Gruyter.

Booth WC (1983) *The rhetoric of fiction*, second edition. University of Chicago Press.

Booth WC (2002) Is there an "implied" author in every film? *College Literature*
29(2): 124-131.

Booth WC (2005) Resurrection of the implied author: why bother? In: Phelan J and
Rabinowitz PJ (eds) *A Companion to narrative theory*. Malden, MA: Blackwell.

Bordwell D (1991) *Making meaning:iInference and rhetoric in the interpretation of
cinema*. Harvard University Press.

Bordwell D (2002) Film futures. *SubStance* 31(1): 88-104.

Buckland W (1995) Relevance and cognition: towards a pragmatics of unreliable
filmic narration. In: Müller JE (ed) *Towards a pragmatics of the audiovisual*.
Volume 2. Münster: Nodus Publikationen.

Buckland W (2009) *Puzzle films: complex storytelling in contemporary cinema*.
Chichester: Blackwell-Wiley.

Currie G (1995) Unreliability refigured: narrative in literature and film. *Journal of
Aesthetics and Art Criticism* 53(1):19–29.

Ferenz V (2005) Fight Clubs, American Psychos and Mementos: The scope of
unreliable narration in film. *New Review of Film and Television Studies* 3(2):
133-159.

Hansen PK (2008) Unreliable narration in cinema: facing the cognitive challenge
arising from Literary Studies. *Amsterdam International Electronic Journal
for Cultural Narratology*, no. 5. Available online at http://cf.hum.uva.nl/
narratology/a09_hansen.htm [view date 19 October 2014]

Kim J (2003) *A Tale of Two Sisters*. Seoul: B.O.M. Film Productions. Bottom of Form

Lothe J (2000) *Narrative in fiction and film: an introduction*. Oxford University
Press.

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 aus Anglistik und
Amerikanistik* 22: 83-105.

Nünning A (1999) Unreliable, compared to what? Towards a cognitive theory of unreliable narration: prolegomina and hypotheses. In: Grünzweig W & Solbach A (ed.) *Grenzüberschreitungen: Narratologie im Kontext.* Tübingen: Gunter Narr.

Nünning, A. (2008) Reconceptualising the history, theory and generic scope of unreliable narration: towards a synthesis of rhetorical and cognitive approaches. In: D'Hoker E & Martens G (eds.) *Narratologia: narrative unreliability in the twentieth-century first-person novel.* Berlin: Walter de Gruyter.

Nünning A (2013) *Unreliable narration : studien zur Theorie und Praxis unglaubwürdigen Erzählens in der englischsprachigen Erzählliteratu.* Trier : Wiss. Verl. Trier.

Olson G (2003) Reconsidering unreliability: fallible and untrustworthy narrators. *Narrative* 11(1):93-109.

Phelan J & Rabinowitz PJ (eds) (2005) *A companion to narrative theory.* Malden, MA: Blackwell.

Phelan J (2007) Estranging unreliability, bonding unreliability, and the ethics of *Lolita. Narrative* 15(2): 222-238.

Sperber D & Wilson D (1986) *Relevance: communication and cognition.* Malden, MA: Blackwell-Wiley.

Wikipedia (2014) Janghwa Hongryeon jeon. Available online at http://en.wikipedia.org/wiki/Janghwa_Hongryeon_jeon [view date 19 October 2014]