

A multimodal analysis of burying strategies used for plot-construction purposes in BBC's *Sherlock*

Kristin Berberich, Nina Dumrukcić
University of Heidelberg

1. Introduction

The purpose of crime fiction is a well concealed mystery. Filmmakers have several tools at their disposal to hide clues and keep audiences guessing. When it comes to television series, an end of the plot might not be in sight, hence filmmakers have to keep the option open for further seasons if ratings hold up or to wrap up the concluding series finale in a satisfying manner. Due to the fact that television series rarely have a linear plot, a macro-context is crucial when determining frequently occurring burying patterns. This presupposes that producers constantly utilise burying processes for information that is yet to be revealed to the viewer in a later plot sequence. Through the usage of flashbacks or other filmic devices, producers can decide to re-iterate elements of previous plot sequences and then foreground or bury different aspects much later, which is why it is important to look at buried instances within the whole context of the series.

This paper seeks to explore at what points in television series clusters of buried items occur and how these are used to keep audiences intrigued for the upcoming season and episodes. In this approach, we look at the macro-context of the series, combining qualitative and partly quantitative methods. The criteria of analysis focus on the linguistic, visual and auditory methods of burying. In our multimodal analysis, we will take into account all currently aired episodes of BBC's *Sherlock*. The musical cues and colour schemes are beyond the scope of this paper and will hence be excluded. The theoretical frameworks against which the analysis shall be carried out is based on Kress and van Leeuwen *Grammar of Visual Design* (2006) as well as Nina Nørgaard's work on multimodality (2010; 2014). Methods of attention manipulation as a means of burying shall also be explored; hereby we draw on Sanford and Emmott's work (2012), which explains how processing capacities and on their cognitive limitations can be utilised by an author, for example through underspecification of an item or relying on the shallow semantic processing.

Clues have to be introduced or hinted at in some way either visually or verbally so the resolution seems credible. Despite this, viewers do not always perceive the clues as what they truly are. This is a careful strategic decision by the makers of a cinematic piece known as

burying. The phenomenon has been studied most notably by stylisticians Catherine Emmott and Marc Alexander (2014) who developed a set of guidelines on how to apply these techniques to crime fiction novels. This paper aims to adapt the model to television series through a case study on BBC's *Sherlock*.

2. Multimodal analysis

Film as a genre is inherently multi-layered. Unlike in literature, the construction of meaning is not limited to a linear horizontal axis realised by the vector direction of the text. Film has the ability to form a complex layering of modes by combining verbal and visual input. Rather than describing consecutively how a character makes an utterance and does something while speaking, the medium of film can present several meaning-making modes at once. Verbal and non-verbal modes (Halliday, 1978) are employed simultaneously in order to create meaning. The challenge for multimodal analysis is to accurately describe instances where several modes work simultaneously in the construction of meaning or, in this case, the construction of plot-sequences and their impact on the viewer's attention with regard to the solution of the case respectively.

Stylisticians who pursue a detailed multimodal analysis that goes beyond the written (McIntyre, 2008; Busse, 2010) have readily taken to the comprehensive toolkit provided by Kress and van Leeuwen. Their *Grammar of Visual Design* (1996; 2006) is based on Hallidayan systemic-functional linguistics and stresses that, even in mundane, everyday interaction, social meaning is realised by a structured 'system' (Kress and van Leeuwen, 2006: 172) that creates social meaning within a text but also with the text's environment. So any multimodal medium contains 'text whose meanings are realized through more than one semiotic code' (ibid., 177). However, their work has also received criticism¹ from film studies scholars, for example because of its occasional reduction of moving pictures to still frames (McIntyre, 2008: 316). Furthermore, film analysts Bateman and Schmidt highlight in their works that

[i]t is often how an indicated shot or shot sequence contrasts with its environment that is significant rather than particular qualities of the shots themselves. Film theory often reminds us of this: there is no 'fixed' meaning for particular camera angles, zooms or other arrangements (Sharff, 1982: 32; Prince, 2007: 22; Bordwell and Thompson, 2008: 192) -

¹ Bateman and Schmidt as well as Bordwell take issue with the term 'grammar' in Kress and van Leeuwen's title and underscore in their works that despite being inherently structured, techniques of filmmaking cannot be compared one-on-one with the grammar of a language (Bateman and Schmidt 2011: 39f, Bordwell 2005: 252).

low angles do not ‘mean’ power, high angles do not mean ‘fate’, and so on. (Bateman and Schmidt, 2011: 17)

Consequently, we cannot disregard the interplay of modes and, perhaps more importantly, the context of a particular filmic decision and its relation to other choices realised in a shot. The context of an entire scene is vital for its development and creation of meaning, and this context can indeed span over the entire episode. Kress and van Leeuwen’s approach is therefore valuable for an analysis of film as long as the issue of context is taken into account. In the analysis of several modes, it is all the more important to find out how these interact with one another to create meaning. In the present paper, the focus will be on the visual and the (non-)linguistic auditory, which are considered the two central *media* in filmic productions by Toolan (2014: 456). Strictly speaking, the two channels are not *modes* themselves since they do in fact serve to express different modes. Language as the linguistic mode, for instance, can be realised in two media, i.e. speech, which is accompanied by many aspects belonging to non-verbal modes, and writing. In brief, ‘modes cut across sensory channels’ (Stöckl, 2004: 11) and act jointly in the creation of a meaning.

3. Burying

In order to keep the plot resolution obscure and the viewers interested, certain elements have to be highlighted so that others can be kept from the viewer. While foregrounded items are perceptually more salient than others, buried items are placed in the background. As is the case in other areas of linguistics, that which is not foregrounded and perceived as highly stylised or frequent has not been studied as extensively (Emmott and Alexander, 2014: 331). In this vein, burying, although being ubiquitously used as a key strategy for attention manipulation in crime fiction, is often merely described as a side-effect of foregrounding. Placing information in the background for plot purposes with the intention of it not being easily found is a highly strategic stylistic choice (ibid., 331). This paper poses the question of how burying is conveyed on television screens. In order for this effect to be achieved, the clues must be carefully concealed within an episode or a series of episodes acting as pieces of a larger puzzle that comes together as a solution, creating an element of surprise or a feeling of satisfaction for viewers who predicted the outcome. These techniques are explored by Emmott and Alexander in several strategies for burying in detective fiction. These strategies have been compiled into a model, as shown in Figure 1.

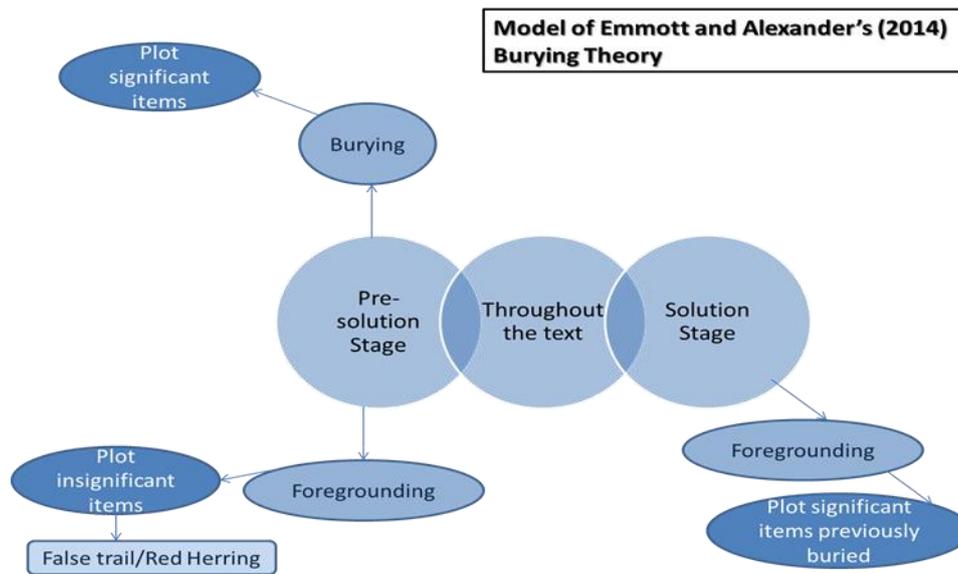


Figure 1.

The strategies are depicted as a model to underline some of the major differences between crime fiction novels and television series. While the former ordinarily have a linear plot, the latter often make use of different modes for the purpose of the storyline. In order to keep the audience guessing, information can be placed at the *pre-solution stage*, occur throughout the entire text or at the *solution stage*, with items being buried or foregrounded accordingly. For these purposes, the term plot is used for the unfolding resolution of the mystery.

If a certain item is heavily foregrounded without contributing to the resolution of the plot, this is called a red herring. The latter are mainly used at the pre-solution stage, or else to foreground aspects that come to be used as red herrings for later (sub) plot sequences that go beyond the scope of the episode in which they are used. At the solution stage, it will finally be specified that this item was not relevant for the plot. False trails are a well-known distraction technique used to create suspense and confusion in the audience (Emmott and Alexander, 2010: 333). In order to provide a credible solution, the clues have to be introduced at some point. An item can be buried at the pre-solution stage while also contributing to the plot resolution at a later stage. These items have to be mentioned in a way that does not draw attention to them, so as not to raise suspicion among the viewers (ibid.).

If items are mentioned at the solution stage, they are previously buried details that are now being highlighted and explained. The unravelling of the mystery should seem plausible in retrospect. This gives the viewer the opportunity to pick up on the buried clues they did not notice whilst watching the series for the first time. The foregrounding effect is achieved by the detective character that puts together the pieces of the puzzle and explains it to the other characters (Emmott and Alexander, 2014: 333). Throughout the text, there must be a moment

of substantial reversal where the red herrings are abandoned and the previously buried items exposed. One example of this can be the emergence of new evidence or a character admitting that they were mistaken. It is also highly recommended to use a combination of these techniques to misdirect the viewer (ibid., 334).

Psycholinguists have suggested that readers are selective about the information they process within a text. By defining how viewers process information in a plot, we can detect when it is less likely for a viewer to perceive a linguistic item and a visual cue as a hint. When the readers of a text pay little attention, information may not be adequately processed. This phenomenon is known as *shallow processing*. Some elements in a text, for instance, may be so familiar and common to the reader that they are perceived automatically and with a lower depth of processing (Sanford and Emmott, 2012: 73; referring to Shklovsky, 1965). One may use this to their advantage in detective fiction, for example by placing a crucial hint within a subordinate clause or amidst other less important items. The suggestion is that not all properties of a word are perceived when it is processed, and either the “lexical or propositional semantic information” (Sanford and Emmott, 2012: 130) is taken into account. By the same token, the processing of film is selective. Filmmakers utilise the fact that “not all aspects of a scene are available to consciousness when a person is presented with a picture” (ibid., 130). Through the complex layering of modes and the high amount of both verbal and visual input, multimodal formats such as films and TV series can hardly be processed to the same degree. Thus, in certain positions in the film, plot-significant elements are less likely to be processed. The depth of processing can vary even more depending on the way these items are portrayed (Emmott and Alexander, 2010: 337). In this line of argumentation, the length of cinematic productions contributes to the fact that the degree of attention fluctuates. Cases of “cognitive misdirection” (Emmott and Alexander, 2010: 328) can be explained by the readers’ varying degrees of attention and their text processing behaviours.

The readers may also be encouraged to focus on the background even though the clue is glaringly obvious. Making heavily foregrounded elements disappear can be accomplished with the help of attention manipulation. Emmott and Alexander state that Agatha Christie, for example, ‘plays on the cognitive limitations of her readers’ (Emmott and Alexander, 2010: 328). Another technique uses a supposedly reliable character to vouch for the reliability of other characters. The readers are more likely to believe information being given by a character that has proven to be trustworthy. The skill lies in making the readers believe false information. Usually the main characters vouch for the reliability of other characters. This, in turn, can lead to the readers wrongly judging characters as reliable (ibid., 333). It is thus

crucial to render the final solution authentic and coherent with the previous text. In order to influence the readers, the author may for instance choose to have the characters mention that the detective's solution is cleverly devised (ibid., 334).

4.1 Visual Burying Through Typography

Typography is not frequently noticed by readers in literary works since it is mostly conventionalised (Nørgaard, 2010: 438). In film, this may be more conspicuous. *Sherlock* is a fine example to be studied because the use of typography reflects the main protagonist's thinking process. The choice of words, but also the size, shape, colour and other characteristics of the font are relevant to the plot and often clues can be buried within the typography. Nørgaard states that 'in visual terms, modality concerns the "truth value" of a given representation- i.e. the question of "as how true" or "as how real" something is represented' (2010: 439).

From the first episode of *Sherlock*, the viewers are informed that visual text represents Sherlock's thinking process, but also media communication. This leads the viewer to pay close attention when typography is foregrounded. The aim of the showrunners is to prove how items can be buried or foregrounded for plot purposes. In the episode *A Scandal in Belgravia*, Irene Adler and Sherlock's first meeting breaks a pattern previously established in the show. When Sherlock meets a new character, nouns, adjectives and phrases are ordinarily portrayed on screen to describe the person and explain Sherlock's deductions. This happens twice in the episode at 00:13:33 when government officials come to escort Sherlock to Buckingham Palace, and again at 00:20:55 with Harry, whom Sherlock had only just met at the Palace. The emphasis here is not specifically on the word choice or semantic meaning, but rather the portrayal of the words and symbols themselves. At 00:25:25, a series of question marks appear, representing Sherlock wondering about the woman.

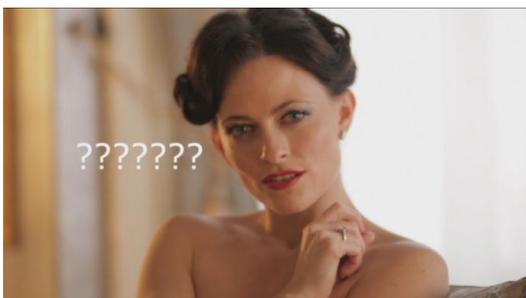


Figure 2.

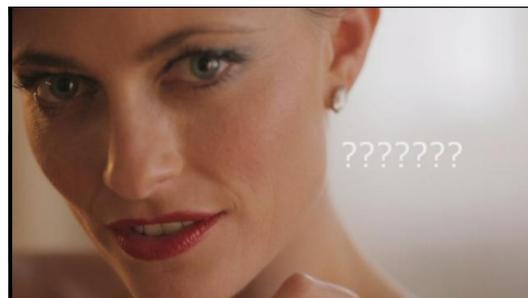


Figure 3.

In the first image of Irene Adler, filmmakers intentionally bury significant traits of her personality, much like Irene herself buries her feelings for Sherlock. Sherlock's ability to read her character and emotions is what leads to her downfall and the mystery being resolved. In comparison, his description of Watson at 00:25:31 is concrete and more elaborate:

(Hasn't phoned sister)



Figure 4.

(Electric not blade)



Figure 5.

The filmmakers wish to stress how much it is bothering Sherlock that he cannot deduce anything about Irene's character from observing her. This is why the size of the font increases when shown alongside Irene. When it is later revealed that Irene's password for her safe was in fact her body measurements, and that Sherlock had noticed this by looking at her body, the meaning behind the question marks could change. He might be wondering what her measurements are. This is an example of foregrounding a plot significant item by focusing only on one feature rather than the main aspect. In this case, one feature was Irene's body, which played a part in the plot for Sherlock obtaining her camera phone, and this was foregrounded by her nudity and the text, but the main aspect, i.e. Irene's emotions which are key to her phone password, is buried as it is not portrayed on screen.

In the episode *His Last Vow*, typography can also be used as red herring. When the character Charles Augustus Magnussen is introduced at 00:24:33 and we see text being shown through his glasses. Here the typography is more similar to a typewriting script used in a previous episode to portray text messaging on a mobile phone or typing on a computer. This leads the audience to believe that he is wearing technologically advanced glasses that show him this information, when in fact it is later revealed to be merely his own thoughts and data he memorised. Regular viewers of the show might arrive at this conclusion because a similar typography was used in another episode when Sherlock is trying to break the password to a mobile phone and the script is more similar to what you would find on an electronic device. The aforementioned examples portray the various ways typography can be used for the purpose of burying information relevant for the plot.



Figure 6.

4.2 Burying within the setting

Another burying technique introduces plot-significant items as a regular part of the setting. Hereby, the attention directed towards these items is reduced. In *A Study in Pink*, this is achieved through numerous descriptions and depictions of a taxi. The viewer is unlikely to assume that this particular means of transport is connected to the case since the taxi is portrayed as ‘heavily schema consistent’ (Emmott et al., 2014: 276). One particular appearance of a black cab at 00:03:40 in *A Study in Pink* serves as an example that, although the scene might be composed to introduce the taxi as a clue to the murder case, the interactive participant, i.e. the viewer, is still processing the scene on a shallow level and hence does not perceive it as such yet. While the two unknown represented participants, i.e. the characters on screen, are walking on the pavement in a dark and rainy night, a taxi approaches from the right hand side. One of the young men says, ‘Yes, yes, taxi, yes!’ and whistles to stop the taxi while waving his arm. At the same time, the sound of the rain and the taxi engine are audible. The date displayed in white font at the bottom centre gives information which is not relevant for the immediate context. The font is inserted to reduce processing of the crucial part of the scene, namely the taxi does not openly appear to be foregrounded as being important for the plot.



Figure 7: Presentation of the taxi as the key information (00:03:47).

This array of linguistic and paralinguistic auditory features alongside with the man's gesture representing the visual level shows that in this sequence of shots, there is more than just one mode at play. At the beginning of the scene, the cab is in the periphery of the shot but the camera perspective soon changes and clearly presents the taxi which drives past the two men. The black cab does not stand out although it is on the right hand side of the frame and construed as the 'key information' (Kress and van Leeuwen, 2006: 180). If we follow Kress and van Leeuwen's framework, the taxi is 'what the reader must pay particular attention to' (ibid.) here. It receives its ultimate significance as a plot-significant item on grounds of its representation in the scene as being the 'new' and not the 'given', as had been the case when Sherlock and John Watson take a taxi from Baker Street which enters the frame from the left.

The fact that the human mind has certain cognitive limitations to it is useful for film producers. In *The Sign of Three*, another plot-significant item is introduced when a catering staff at John and Mary's wedding pulls a meat poker, which closely resembles the murder weapon used to kill the guardsman, out of a roast. Through an interplay of various modes, the item is buried in its immediate context by reducing its prominence and importance in the scene. This is perhaps even more striking, since this item is not connected to the main characters of the show in any way, such as for instance the taxi, and is thus not rendered credible just by being associated with them. In the scene, the man who handles the meat poker and not the actual item that is used to commit murder is the focus of the conversation camera. The item is not present on the linguistic mode, while a great amount of visual input is given. On the visual level, only one shot of the poker is given while various other attributes of the man, such as his smart phone, complexion, and character traits, are foregrounded. These

elements are taken up by the represented participants Sherlock and Janine on the linguistic auditory level and focussed on by the camera. In this scene the context of the shot is spiked with plot-insignificant details which are presented in a very close shot. This happens in such a quick succession that the viewer might have trouble processing every single piece of information, but especially the information that is not foregrounded by the camera. Shallow processing thus enables the filmmakers to show and subsequently bury the meat poker in the immediate context since the attention is directed elsewhere.

In *The Empty Hearse*, passenger transport on rail has been established as a crucial part of the city by John Watson's journey on the tube and a series of visual cues. Here, related word senses are to create ambiguity as an attempt to cover up the significance of a plot-significant item. Sherlock gives a hint towards the nature of the terror threat when he says in 00:13:45, 'I will find your underground terrorist cell, Mycroft.' (Sherlock, 3.1), According to the OED, an extended use of *cell* is 'compartment; an enclosed space [...]' (OED: 'cell, n. '), which clearly relates to the semantic field of transport on rail. Furthermore, Sherlock is called to a crime scene and in the midst of his observations, at 00:42:22, dust trickles off the ceiling and a distant noise can be heard. Molly Hooper, who functions as his assistant in Watson's place at the time, asks him about the noise, 'Trains?' and Sherlock simply states, 'Trains.' (Sherlock, 3.1) This would not be of interest if it did not contribute to the creation of the semantic field and if it was not completely random in the context. The description of the train noises is underspecified and is not given any value in the fictional film world¹. The plot-significant information is subsequently buried by Sherlock's analysis of the corpse with a digital insertion of a compass into the filmic image, which is done to illustrate Sherlock's thought process, and a camera zoom on parts of the corpse's clothing. This shows that the detective is preoccupied with the present case. In this context, Sherlock is also rendered unreliable as a character, because he seems to hear his former partner Watson's voice, which is put into words in typography for the interactive participants, in his head. Portraying the consultant detective thus makes him seem unreliable and consequently buries the repeated linguistic item 'train' with the help of contextual filmic decisions. In conclusion, the viewer is misdirected although presented with all the clues that are needed to solve the crime plot. This is achieved by using linguistic items from a semantic field with related senses which cause ambiguity and help to cloud the perception of the information given.

5. Conclusion

Burying strategies employ various modes at a time to introduce plot-significant items without directing attention to them. In our analysis, we found that an array of linguistic, non-linguistic auditory, as well as visual features is used. Furthermore, items are buried across the macro-context of the series, with specific items significant for the plot clustering around certain plot stages. Items which are buried at the beginning of an episode are relevant for the main crime plot, while items buried in the throughout-the-text stage seem to have relevance for future (sub) plots as they are presented too close to the solution of the crime to be buried credibly.

Far from being irrelevant in terms of burying, the solution stage near the end of an episode can indeed be utilised to bury items or clues whose relevance will become more discernible in future episodes. We argue that these are not simply foregrounded but buried by their immediate context because of the compositional choices made by the producers. This can work since there is not such a rigid end to the plot in a TV series as for example in a novel or movie. On the contrary, filmmakers do introduce plot-significant items at these seemingly ill-fitting stages of the plot to keep audiences intrigued and willing to follow the series. The appearance of said clusters can be visualised for the whole entire series and we suggest that there will be visual clusters with a high burying density (on several modes) which serve different functions in terms of plot-construction. The stylistic study of television series can offer insights into complex webs of meaning-making processes spread out over several layers from the visual to the aural. But due to the specificity of the genre which invites the viewer to follow the represented participants over large time-spans in the film world, the level of burying density and the sheer width of the time span during which items can be buried invites us as stylisticians into new depths in the analysis of burying in televised crime fiction.

Bibliography

- Bateman, John and Karl-Heinrich Schmidt (2011) *How Films Mean*. London: Routledge.
- Busse, Beatrix (2010) 'Non-Literary Language: A Stylistic Investigation of the Cover Pages of the British Satirical Magazine *Private Eye*', in Dan McIntyre and Beatrix Busse (eds) *Language and Style*, pp 468-500. London: Palgrave Macmillan.
- Emmott, Cathy and Marc Alexander (2010) 'Detective Fiction, Plot Construction, and Reader Manipulation: Rhetorical Control and Cognitive Misdirection in Agatha Christie's *Sparkling Cyanide*', in Dan McIntyre and Beatrix Busse (eds) *Language and Style*, pp 328-346. Basingstoke : Palgrave Macmillan.
- Emmott, Catherine, and Marc Alexander (2014) 'Foregrounding, burying and plot construction', in Peter Stockwell and Sara Whiteley (eds) *The Cambridge Handbook of Stylistics*, pp 329-343. Cambridge: Cambridge University Press.
- Gavins, Joanna (2007) *Text world theory: an introduction*. Edinburgh: Edinburgh University Press.
- Halliday, M. A. K. (1978) *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- Kress, Gunther and Theo van Leeuwen (2006) *Reading images – The Grammar of Visual Design*. London: Routledge.
- Kress, Gunther and Theo van Leeuwen (2001) *Multimodal Discourse: The modes and media of contemporary communication*. London: Edward Arnold.
- McIntyre, Dan (2008) 'Integrating multimodal analysis and the stylistic of drama: a multimodal perspective on Ian McKellen's Richard III', *Language and Literature* 17(4): 309-334.
- Nørgaard, Nina (2010) 'Multimodality: Extending the stylistic tool kit', in Dan McIntyre and Beatrix Busse (eds) *Language and Style*, pp 433-448. Basingstoke: Palgrave Macmillan.
- Nørgaard, Nina (2014) 'Multimodality and stylistics', in Michael Burke (ed) *The Routledge Handbook of Stylistics*, pp 471-484. London: Routledge.
- OED Online:
"cell, n.1". OED Online. December 2014. Oxford University Press.
<<http://www.oed.com.ubproxy.ub.uni-heidelberg.de/view/Entry/29468?rskey=ZySHdU&result=1&isAdvanced=false>>, accessed 25.11.15.
- Sanford, A.J. and Catherine Emmott (2012) *Mind, brain and narrative*. Cambridge : Cambridge University Press.
- Sherlock. Dir. Paul McGuigan and Euros Lyn. Perf. Benedict Cumberbatch, Martin Freeman. 25 July 2010. DVD. Hartswood Films/BBC Wales/WGBH, 30 August 2010.
- Stöckl, Hartmut (2014) 'In between Modes: Language and Image in Printed Media', in Eija Ventola, Cassily Charles and Martin Kaltenbacher (eds.) *Perspectives on Multimodality*, pp 9-30. Amsterdam: John Benjamins Publishing.
- Toolan, Michael (2014) 'Stylistics and film', in Michael Burke (ed) *The Routledge Handbook of Stylistics*, pp 455-470. London: Routledge.
- Werth, Paul (1999) *Text Worlds: Representing Conceptual Space in Discourse*. New York: Longman.

¹ The idea that a given text influences its own reception and enables the viewer to construct a text world is formulated for literary texts in text-world-theory (cf. Werth 1999, Gavins 2007).