Multimodality and audiovisual translation

Cohesion in accessible films

Aline Remael and Nina Reviers

Introduction

Multimodality is fast becoming the main conceptual framework for the study of audiovisual texts, i.e. texts that create meaning through the use of multiple semiotic modes, such as films. Individual semiotic modes include the aural-verbal mode (dialogues and lyrics); the aural, non-verbal mode (music and sound effects); the visual-verbal mode (various types of text on screen); and the visual, non-verbal mode (images) (Delabastita 1989, Remael 2001, Zabalbeascoa 2008). Multimodality examines how these individual modes function and how they can be combined into a unified whole. All the modes have a role to play in the creation of meaning in a text, but their importance will vary: in some films or film scenes music may be dominant; in others, the images may carry the story forward. What is more, combining different modes creates supplementary meaning, on top of the meanings conveyed by the individual modes. This is what Baldry and Thibault (2006) refer to as the ‘resource integration principle’. What matters, however, is ‘how they [the semiotic modes] all add up and combine with each other so that viewers can interpret them in certain ways’ (Zabalbeascoa 2008: 25). To guide users’ interpretations, filmmakers insert different types of implicit and explicit links between modes, which serve as cues for the users to reconstruct a coherent end product. In this chapter, multimodal texts are seen to work through multimodal cohesion, which is accomplished through different cross-modal ties between the different semiotic modes. We pragmatically define multimodal cohesion as any instance of implicit or explicit ‘sense-relation’ between two or more signs, from the same or different modes, within a given text that helps the viewer create a coherent textual semantic unit (for an overview of key concepts in multimodal theory, see the publications listed in the Further Reading section, at the end of this chapter).

Multimodality and multimodal cohesion in audiovisual texts remain complex issues that have been addressed from different angles. The challenge of multimodal research today involves designing a systematic framework for the analysis of multimodal texts and their cohesion, especially in the context of film and film translation. To date, three different approaches have been deployed to do so. In Introducing Social Semiotics (2005: 179), van Leeuwen approaches multimodality from the perspective of social semiotics and puts
forward four concepts that guide the integration and co-occurrence of different kinds of semiotic resources: rhythm, composition, information linking and dialogue. Rhythm is seen to provide coherence and meaningful structure to events that develop over time, whereas composition does the same within a spatial dimension. The concept of information linking is used to refer to the cognitive connections between items of information in time and to refer to the temporal or causal links between words and images in space-based multimodal texts. Finally, for van Leeuwen, dialogue is a broad term that encompasses the structures of dialogic exchanges (i.e. not only human dialogue) and musical interaction in multimodal texts and communicative events. However, these terms remain open to interpretation when applied in different contexts, as do the relations between them. As formulated by van Leeuwen, this set of concepts does not constitute, nor is it meant to be regarded as a systematic analytical framework.

Many scholars have built on van Leeuwen’s pioneering work trying to come up with a more specific conceptualization of co-occurrence and cohesion across different semiotic signs. One of them is Royce (2007), who accounts for the co-occurrence of different semiotic modes in terms of the linguistic concept of lexical cohesion. From this standpoint, the integration across modes is based on reference, repetition, synonymy, collocation and part-whole relations—including meronymy (the semantic relationship that holds between a car steering wheel and the car) and hyponymy (the semantic relationship between a spoon and the wider category of cutlery that includes it). Royce uses these linguistic concepts to name the cohesive ties that exist between non-verbal modes. For example, a shot of a woman’s hand writing a letter prompts the viewers to link the hand to a woman they have seen in a previous scene (The Hours 2002)—in what could be analyzed as an example of meronymic relationship.

However, we have found that these categories do not necessarily cover all instances of co-occurrence in filmic texts, because some instances are driven primarily by implication (based on the mental models that viewers bring to their spectatorial experience) and dialogic interaction (comparable to van Leeuwen’s concept of dialogue). ‘Complementarity’ is thus suggested as an additional concept that builds on van Leeuwen’s concept of information linking. It is meant to designate the relationship between two signs that simply co-occur or appear in each other’s immediate textual environment. For example, a man interrupting his telephone conversation and saying ‘hold on’, as a reaction to mumbling and shouting in the streets outside his office (Süskind 2012). This concept can enable subtle distinctions, as the difference between lexical cohesion and complementarity can be a matter of degree (see Reviers and Remael 2015).

Chiao-I Tseng (2013) has recently made an interesting contribution to this discussion. Her work builds on the social semiotics tradition and makes an initial attempt to bring together the above-mentioned concepts under a systematic analytical model. Tseng (2013: 153) states that ‘viewer’s comprehension of film narrative is premised on the meaning patterns they have constructed somewhere else, either from their previous film viewing experiences or from other life experiences’. In this sense her approach is akin to the theory of mental models developed by Johnson-Laird (1983) and regularly applied in narratology, which contends that readers and viewers processing narratives construct mental representations of such narratives (see, for instance, Herman 2009). Similarly, Tseng (2013: 1) demonstrates that film viewers attend to four elements to construct a coherent film narrative—characters, objects, settings and action—that she uses as her starting point to analyze filmic cohesion. Based on insights from social semiotics and systemic functional linguistics, she argues that ‘the structure of identification, namely, how relevant people, places and things are actually tracked,
highlight the unity of a particular text’ (ibid.: 39). In other words, viewers track the identities of characters and objects by constructing cohesive chains that guide their narrative interpretation. Tseng’s approach is particularly relevant to audiovisual translations (AVTs). Our own, admittedly limited, research on cohesion in AVT (Reviers and Remael 2015) suggests that cohesion is not only created through the use of, for instance, explicit references and lexical cohesion, but also by more implicit sense-relations that become apparent by tracking characters, their actions and behaviour. In short, the ease with which the basic, narratological building blocks (characters, setting and time) can be identified by users across audiovisual texts seems to be a crucial manifestation of cohesion between such texts and their translations, and is therefore an appropriate starting point for the present study.

**Research problem**

Multimodal cohesion constitutes a particular challenge in AVT and media accessibility. In AVT one mode is altered/translated and, as a result, the explicit or implicit interaction between the translated mode and the other modes may also be altered, sometimes unintentionally. Translators of audiovisual texts must therefore be cautious not to break down the multimodal cohesion of the source text and recreate a coherent and cohesive end product. This is particularly true for assistive forms of audiovisual translation such as audio description (AD) and subtitling for the deaf and hard of hearing (SDH).

The aim of the present chapter is to discuss and illustrate how multimodal cohesion is maintained or (re)created in an accessible film clip with AD for the blind and visually impaired and SDH using Tseng’s model. AD makes audiovisual products accessible for blind and visually impaired users by translating the visual mode of, for instance, fiction films (this chapter’s topic) into an aural-verbal narration that makes use of the original verbal-aural and non-verbal-aural modes of the film—i.e. its dialogues, sound effects and music—to produce a new coherent, purely aural target text. SDH translates the aural-verbal dialogues as well as narratively relevant sounds and music into subtitles, a visual-verbal mode that also involves the use of colours or other methods for speaker identification. In both assistive forms of AVT, one mode is entirely replaced by another and the end product should also function mono-modally for the target audience (visually in the case of SDH and aurally in the case of AD). The analysis presented below pays particular attention to the role of sound in AD and SDH, as this aspect has not received sufficient attention to that, even though several authors have acknowledged the importance of sound in both assistive modalities. Neves (2010), for instance, tackles the challenge of incorporating music in SDH, proposing strategies to move beyond the current verbal and symbolic renderings typically used in SDH; Igareda (2012), on the other hand, illustrates the importance of music and lyrics in audio described films. For her part, Fryer (2010) provides valuable insights into the importance and functions of sound in AD.

**Analysis**

This section delivers a detailed analysis of a short clip from *Nights in Rodanthe* (Wolfe 2008). The film tells the story of Dr Paul Flanner, a (plastic) surgeon who retreats to the tiny coastal town of Rodanthe in the Outer Banks of North Carolina, where he stays at an inn run by Adrienne Willis. As Paul arrives, a major storm is forecast. With the storm closing in, they turn to each other for comfort and begin an intense romance. The selected scene features Paul, just after his arrival at the inn, jogging along the beach, thinking back
to a surgery he performed a few weeks earlier. *Nights in Rodanthe* was selected because it was available with both AD and SDH. This specific scene was chosen because it stands on its own as a coherent unit, features all semiotic modes and has important sound and music elements, which are bound to be impacted in the accessible versions. Table 17.1 features a multimodal transcription of the clip.

**Methodology**

The social semiotic strand of multimodal theory, and in particular the work of van Leeuwen, Baldry and Thibault, has developed a method for film analysis known as multimodal transcription. Together with the development of multimodal corpora, this sort of transcription has become a much-used methodological tool in multimodal research (see e.g. Pérez-González 2014: 165; 295 and Baldry and Thibault 2006). Multimodal transcription is a method that supports the detailed and empirical analysis of the functioning and interplay of all the components of audiovisual texts. This methodology, initially developed for film analysis, has been applied to the analysis of translated film by several authors, including Taylor (2004a, 2004b). The idea is to segment the text into its smallest constitutive units in order to facilitate an objective and empirical analysis. In brief, the transcription consists of a column-based table, in which the visual and auditory components of the film clip are transcribed on the basis of logical, semantic units. Significantly, ‘[t]he number and ordering of the columns included in the transcription, the range of modes and sub-modes covered in the analysis, and the set of notation conventions used for coding purposes depend on the specific needs of the individual project’ (Pérez-González 2014: 295). The transcription for the purpose of this study is available in Table 17.1, and consists of six columns: unit number, transcription of the AD, coding of the visual image and kinesic action, transcription of the sound and music, transcription of the dialogue and, finally, transcription of the SDH. Each unit largely corresponds to one shot delineated by a clear cut. Items listed in the same row occur simultaneously. The length of cells within columns signals continuity.

**Theoretical framework**

As indicated above, Tseng focuses on the process of identification that filmic texts use to ensure that viewers can track the identity of participants (i.e. characters, settings and objects) and their cumulative development, across sequences. The framework for the analysis of filmic cohesion that she develops contains four analytical tools, which constitute the basic building blocks of our analysis: the presenting system, the presuming system, the comparative system and the system for the direction of identity retrieval.

Characters can be presented on the basis of three parameters. First, they can be presented as generic (e.g. as a teacher, member of a larger group of teachers) or as specific individuals (e.g. a teacher who is named or known). However, the distinction must be seen as a continuum that is subject to change, as generic characters can evolve into named individuals while the narrative unfolds. Second, both a generic and a specific presentation can be realized mono-modally or cross-modally, by two or more modes simultaneously. Third, Tseng’s framework is based on the analysis of the most salient elements of the narrative. Participants can acquire salience immediately or gradually. In the case of a gradual presentation process, the options are either a dynamic or a static method. In a dynamic introduction a participant becomes more salient gradually, e.g. through different shot types. In the case of a gradual but static introduction, the film makes use of a prelude (for instance, the sound of a car is heard,
### Table 17.1 Multimodal transcription

<table>
<thead>
<tr>
<th>No.</th>
<th>Audio Description</th>
<th>Visual Image + Kinesic Action</th>
<th>Sound + Music</th>
<th>Dialogue</th>
<th>SDH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[FEMALE VOICE THROUGHOUT] Dawn. Dressed in a hooded sweat top, Paul descends the steps from his . . .</td>
<td>Horizontal, median, stationary shot from the back end of a wooden balcony overlooking the ocean on the horizon. Paul, in jogging suit, slowly walks from the left of the screen to the edge of the balcony in the middle of the frame, with his back to the camera (from close to medium shot distance). Low-key lighting and grey colours.</td>
<td>Footsteps on wooden stairs and slow ocean waves in the background.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>. . . balcony to the beach. He begins to jog, then race along it. He walks along a hospital corridor accompanied . . .</td>
<td>Cut to a tracking shot starting underneath the balcony stairs and moving right towards the beach. The visual focus moves from Paul’s feet on the stairs (close distance), to Paul in the centre of the image on the beach running (long shot distance). Tempo of the kinesic action moves from slow walking to running, then racing. Low-key lighting and grey colouring continue.</td>
<td>Slow ocean waves and footsteps on the sand, gradually going faster. As the footsteps go faster, up-tempo piano music joins in (building up from low to high-pitched notes). Ocean waves gradually make place for rhythmic, fast, loud breathing.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>. . . by a team of people.</td>
<td>Cut to a medium (tracking) shot of Paul’s face and upper body. He is gazing into the distance along the beach. Low-key lighting and colouring continue. The tempo of the kinesic action remains high.</td>
<td>Breathing and up-tempo piano music continue.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Cut to a close-up of Paul’s running feet on the beach (high-angle, tracking shot). Low-key lighting and colouring continue. Tempo kinesic action remains high.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>Cut to a close-up of Paul’s face and upper body, against the ocean in the background. (Identical to shot 3).</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6 Cut to a low-angle, tracking shot of Paul walking quickly in doctor’s suit along a brightly lit (high-key lighting), brightly coloured hospital corridor, surrounded by nurses. Paul is filmed from the front, reading a chart a nurse is holding in front of him. Another doctor cuts across the screen to quickly shake Paul’s hand, then walks off screen.

Up-tempo piano music continues.

[PAUL’S VOICE]:
Any problems with anaesthesia?

[NURSE’S VOICE]: no problems.

[PAUL’S VOICE]: absolutely not, you’re sure?

[DOCTOR’S VOICE]: Paul. [PAUL’S VOICE]: Hi John.

— Any problems with anaesthesia?
WOMAN: No problems.

— Absolutely none? You sure?
WOMAN: yes.

7 Cut to a close-up shot of Paul’s face and upper body as he is running along the beach (identical to shot 3 and 5).

Up-tempo piano continues, fast, rhythmic breathing joins in again.

8 He’s in . . .

Cut to Paul in the hospital corridor: low-angle, tracking shot, following Paul (right) turning left towards a nurse and walking past her, disappearing through electronic doors opening on the far left of the screen, followed by the group of nurses. On the door is written: ‘operating room 1’. High-key lighting, bright colours and high tempo kinesic action.

Up-tempo piano music continues.

[NURSE’S VOICE]: Miles Davis again, doctor Flanner? [PAUL’S VOICE]: Bach, today, Bach!

— Miles Davis again, Dr. Flanner?
— Bach today.
Bach.

9 . . . theatre, preparing a patient’s face for surgery.

Cut to tracking shot of Paul on the beach again; identical close-up of his face and upper body (shot 3, 5 and 7).

Up-tempo piano continues, fast, rhythmic breathing joins in again.

10 Stationary, slightly low-angled shot of Paul (median close-up) in the middle of the frame, in the operating room in scrubs. Contrast lighting created by a bright overhead light in the top right corner of the image.

Up-tempo piano music continues.

(continued)
Table 17.1 (continued)

<table>
<thead>
<tr>
<th>No.</th>
<th>Audio Description</th>
<th>Visual Image + Kinesic Action</th>
<th>Sound + Music</th>
<th>Dialogue</th>
<th>SDH</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Bright overhead . . .</td>
<td>Cut to a low-angle shot of Paul in the operating room in his scrubs. Visual focus on Paul's hand in close-up at the bottom of the screen, his head right above and four bright lights surrounding his head, creating contrast lighting. The tempo of the kinesic action (hands moving) is high.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>... lights blaze . . .</td>
<td>Cut to a top-shot of a patient lying on Paul's table. Visual focus on the patient's face lit by a bright light, surrounded by hands, green cloth, and surgical material in darker light. Kinesic tempo of the hands remains high.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>. . . behind him.</td>
<td>Cut to a close-up, top-shot of the patient's brightly lit face, half covered by green sheets and two hands measuring a scar on the patient's face, and giving an injection. Kinesic action remains high.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Paul races along the beach. The patient's . . .</td>
<td>Cut to a low-angle shot of Paul in scrubs, surrounded by four bright overhead lights (contrast lighting continues). As he bends forward, his headlight blinds the camera. Kinesic action of hand movements is median. Up-tempo piano music continues and when Paul bends towards the camera, ocean waves mix in.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>. . . heartbeat flat lines.</td>
<td>Cut to a median close, tracking shot of Paul's face and upper body again on the beach (identical to shots 3, 5, 7 and 9) Up-tempo piano music continues, ocean waves become louder and rhythmic breathing joins in.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>Cut to a low-angle, medium close-up of Paul in scrubs in the operating room, looking up alarmed (gazing at something off-screen to the left). Contrast light created by a big overhead light next to Paul's head. Loud, high-pitched beeping and low (unclear) voices. Up-tempo piano continues. [EKG BEEPING RAPIDLY]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>As Paul further turns his head, a cut follows to a shot of a heart monitor and other screens behind Paul. The camera quickly zooms in on the flat line on the heart monitor. Contrast lighting continues. Loud, high-pitched beeping and muffled voices. Up-tempo piano continues.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
18 Cut to a high-angle, overview shot of the operating room. Visual focus is on the brightly lit patient in the middle, surrounded by doctors, nurses, and the equipment in darker shadows. Beeping sound of charger and beeping of heart monitor. Up-tempo piano continues. [FEMALE VOICE]: the patient’s gone. [MALE VOICE]: charging. WOMAN: charge. MAN: Charging.

19 Cut to an extreme close-up of the heart monitor. Loud beeping and up-tempo piano music continue. [PAUL’S VOICE]: Talk to me . . . Talk to me! Come on, now!

20 Cut to the close-up of Paul’s face on the beach (identical shot 3, 5, 7, 9 and 15) [PAUL’S VOICE]: . . . . . come on now. Three, four . . .

21 Cut to a close shot of Paul in the operating room, looking down. Three, four five, six . . .

22 Cut to a high-angle, medium close shot of Paul in his scrubs performing heart massage. Paul is high up in the frame, the patient’s brightly lit face is visible in the lower left corner. Kinesic tempo is high. [PAUL’S VOICE]: . . . five, six.

23 Cut to a medium, horizontal shot of a man standing behind the operating room door. Contrasting light turns him into a dark silhouette. [PAUL’S VOICE]: . . . seven, eight, nine, 10, 11. . .

24 As Paul . . . Cut to a medium-close, horizontal shot of Paul doing heart massage in his scrubs (identical shot to 21) Loud beeping and up-tempo piano music continue. Soft voice. . . seven, eight, nine, 10, 11. . .

25 . . . pumps his patient’s chest . . . Cut to a close-up of the man’s face behind the operating room door. Soft voice continues.

26 . . . he looks over his shoulder Cut to a horizontal, medium shot of Paul doing heart massage in his scrubs (identical shot to 21 and 24). He looks over his shoulder to the right, while continuing heart massage. Soft voice continues.

27 . . . and sees Mark watching anxiously. Cut to a close-up of the man’s face behind the operating room door, slowly looking down. Soft voice continues.

Legend: bold = character identification chain| underlined = setting(s) identification chain| italics = present time period | regular print = flashbacks.
and then the car appears). Again, all these operations can be performed mono-modally or through a combination of different film modes. In the clip under analysis, the most salient element moving the story forward is the protagonist. Different camera techniques coded in the transcription gradually present Dr Flanner as the most salient element and he remains central throughout the clip (see, for instance, rows 1 to 3 in Table 17.1). Our analysis will therefore focus on the identity tracking of Dr Flanner. Next it will examine the tracking of the settings and a few specific objects in them that are closely associated with the character.

The second notion of cohesion in the framework is that of identity tracking through the presuming system (ibid.: 60). Tseng explains that ‘the presuming systems set out features that bring about the re-identification and tracking of filmic participants after [our emphasis] they are presented in a film’. Characters, settings and objects are presented gradually throughout the film and each scene contributes new elements to the previously constructed mental model of that participant. Characters, objects or settings are not presented only once; instead, they normally reappear. As our analysis shows, however, the processes of presentation and reappearance (presuming) of participants often overlap. Elements can reappear explicitly (either mono- or cross-modally) or implicitly, e.g. when some physical parts of a visual participant reappear in different shots.

Next, there is the comparative system. Filmic comparative elements can indicate a relation of similarity or difference between two shots of a participant, through attributes of quantity (generic participants presented in groups or as individuals) or of quality, such as tempo and volume in the audio mode or colour and lighting in the visual mode. There is an obvious connection here with van Leeuwen’s concepts of rhythm and composition.

The final step in the framework is determining the direction of the identity retrieval. The identity of a participant can be determined endophorically (from the text itself) or exophorically (from the context of situation/culture). According to Tseng, cohesion is normally concerned with endophora only. But while endophoric links are primarily derived from within the text, they are also articulated by mental models that rely both on the viewing experience and the viewers’ knowledge of the world. For Tseng such endophoric cohesion includes ‘indirectly retrievable’ information, based on implicational relations—more specifically lexical relations such as synonymy, hyponymy and meronymy, similar to Royce’s interpretation of intersemiotic complementarity presented earlier. This type of cohesion, that she calls ‘bridging’ (ibid.: 70), occurs when the link needs to be inferred. Bridging, which can be cued mono-modally or cross-modally, resembles complementarity in our original terminology, as it designates instances where two signs complement each other semantically simply by co-occurring in each other’s immediate textual environment. When analyzing links between non-verbal filmic modes, cohesive relations based on juxtaposition are achieved through film editing and it is the viewer who retrieves the relevant links by implication.

**Presenting and tracking characters and settings**

We now turn to our clip to identify the consecutive semiotic signs that trace the identification of the character and the setting, thereby forming a cohesive chain. For details relating to the multimodal narrative into which the chains fit as well as details about camera positions and film techniques, readers are referred to the numbered units in the transcription available in Table 17.1. The clip consists of two interrelated scenes, which also represent two sub-chains insofar as the identity tracking of participants is concerned: one that tracks participants in a past time period and setting, and one that tracks participants in a present time period and
setting. After analyzing the source text in detail, we examine the same identity chains in the accessible versions of the film with AD and SDH.

**Tracking the character in the original clip**

Dr Paul Flanner, the film protagonist, is known to the viewers because he has already been presented. Therefore the scene builds on the mental model the audience has of Paul as a person, meaning he is depicted here as a specific (not a generic) person. Hence the discussion below focuses on the ‘presuming system’—i.e. the reappearances of Paul that together form a cohesive chain—which allows viewers to keep identifying Paul as Paul and the ‘comparative system’ to indicate similarity or differences between the ways in which he reappears. Overall, Paul’s identity is tracked cross-modally. The input of the visual mode is predominant and is supported by two aural modes: sound and minimal dialogue.

In **unit 1** Paul makes an explicit visual reappearance as a specific, known character: he can be seen walking into the scene and ending up in the middle of the frame. He is thereby identified as salient. Like in all the following units, this reappearance works anaphorically, as it points backwards to the referent that was introduced earlier in the film. His identity can be retrieved directly, as we can see his whole figure and directly recognize him as Paul. His reappearance is expressed cross-modally since it is accompanied by the sound of footsteps on a wooden balcony. **Unit 2** features a shot change to a close-up of Paul’s feet. His character makes an implicit visual reappearance, accompanied by an explicit acoustic reappearance of the sound of footsteps, and audiences have to infer that the feet refer to the same referent by an anaphoric, meronymic cohesive relation (i.e. bridging). In the same unit, there is a second reappearance of Paul, when his whole figure gradually becomes visible while he walks further away from the camera and starts running. Here he makes an explicit visual reappearance, as he can be directly recognized as the same Paul we saw in unit 1. His identity is tracked cross-modally again, as the sound of footsteps reappears, this time implicitly; in other words, the sound of the footsteps refers to the same referent, but the audio quality (comparative system) is different, as the feet now walk on sand. Simultaneously, another cross-modal identification cue is added: loud breathing. The next three alternating units in the character identification chain are all linked anaphorically and meronymically (implicit reappearance through bridging): a close-up of Paul’s face (**unit 3**), one of his running feet (**unit 4**) and another close-up of his face, identical to the previous one (**unit 5**), all simultaneously cued aurally as well by the sound of footsteps and breathing. It must be noted that the identification of Paul in units 1, 2, 3, 4 and 5 is supported by the comparative system, as there is a clear similarity between the reappearances expressed through quality attributes in the different modes: Paul is wearing the same jogging-outfit in all implicit and explicit reappearances and low-key lighting, colour and mise-en-scène are the same in the different units. The music in these units also contributes to the relation of similarity between the reappearances, more specifically through the system of quality (non-diegetic): the rhythm and tempo of Paul running along the beach (footsteps and breathing) are mirrored in the rhythm and tempo of up-tempo piano music, creating a strong cohesive relation between the different visual and aural modes in this sequence.

In **unit 6** the camera gradually moves from a long shot of Paul in a white doctor’s coat to a medium shot with a frontal view of Paul down to his waist. The character makes an explicit, visual reappearance, and can be directly identified. The identification is again cross-modal, as dialogue kicks in with a voice linked anaphorically and aurally to Paul through previous scenes of the film, but also linked to him through the visual mode—since dialogue and the
visual sign of Paul in doctor’s coat coincide and are linked to him through the content of the dialogue (they are discussing an upcoming surgery). However, the comparative mode here suggests that the Paul identified in this specific unit occurs in a different time period, as he is presented as a different person through the quality attributes of clothing, i.e. a doctor’s coat (diegetic) and through the different, high-key lighting and bright colours in the hospital corridor (partly diegetic, partly non-diegetic). Another important cross-modal cohesive element in this unit that is expressed through the comparative system is the music: the up-tempo piano music resembles the fast tempo of the kinesic action (fast walking) and of the dialogue (fast talking). What is more, it is the same piano music as in the previous unit where Paul was on the beach (sound-bridge), creating a cohesive link with earlier reappearances of Paul in the identity chain. As one can see in the transcription, this cohesive role of the music continues along the whole identity chain.

In unit 7 we return to Paul running on the beach. He makes an explicit, directly retrievable reappearance that is tracked cross-modally: we see a close-up of his face and we hear his breathing. These cues have a clear filmic similarity (comparison) to the previous elements in the chain, as their quality is the same (lighting, colour, tempo, rhythm) as in units 1 to 5. At the same time they create a relation of difference with the directly preceding cue in unit 6, where lighting, colour, tempo and clothing of the character were different. The continuing piano music again creates a cohesive link of similarity across all the chain elements.

Unit 8 cuts back to Paul at the hospital with an explicit visual reappearance that is directly retrievable. The relation to previous elements in the identity chain is again cued through the comparative system (lighting, clothes, rhythm, tempo). Paul’s identification is supported here by the aural mode, as we can hear his voice, and a cue from the dialogue, as a secondary character calls him by his name. In unit 9 we briefly return to Paul on the beach, where he is identified cross-modally in the same way as in unit 7.

Units 10–14 depict Paul in the operation room preparing a patient for surgery. The tracking of Paul is achieved visually here, through a quick follow-up of implicit and explicit visual appearances: we see Paul’s face, upper body and then his hands (which are linked through meronymy) and we see Paul wearing green scrubs (which can be linked to the cueing of Paul in white doctor’s coat earlier through hyponymy since a white doctor’s coat and green scrubs are both part of the larger category ‘clothes hospital personnel wear’). What stands out when we look at the non-verbal cueing in this series of units is that there are no sound-effects (only the piano music continues) until unit 14 when the sound of waves is heard faintly in the background—even though visually we are still in the operating room. Even though the sound of waves has no co-referential relation to Paul’s identity, it can be seen as part of his identity chain, as it serves as a leitmotif on the one hand (the sound of waves was heard earlier in unit 1 when we saw Paul on the beach and is possibly linked in the mental model of the audience to the identity of the jogging Paul), or it can be seen as a prelude, since in the subsequent unit (unit 15) this recognizable sound of waves becomes louder and is combined with a visual shot of Paul on the beach, tracked in the same cross-modal way as in units 7 and 9.

Units 16 to 27 depict Paul during surgery in the OR, intertwined with one flashback to the same beach scene (unit 20). Paul in the hospital is cued by a series of explicit, visual reappearances that are directly retrievable of his face and upper body in scrubs (units 16, 17, 20, 21, 22, 24 and 26). The comparative system creates a relation of similarity with previous appearances in the operating room and a relation of difference with the identity cues of Paul on the beach. One cue of the identity chain in this sequence stands out: in unit 18 a top-shot of the operating room depicts a patient in the middle of the frame surrounded by a group of
doctors and nurses in similar green scrubs. There is a co-referential link in this shot to Paul’s identity chain that is indirect (bridging) and can be established through hyponymy: in the mental model of the audience, Paul has been identified as an individual, being part of the larger generic group of ‘doctors and nurses’. This identification through hyponymy is supported by the comparative system: there is a similar relation (through the quality system), namely the green scrubs they all wear, but also a difference relation in terms of the quantity system, Paul is now seen as part of a larger group.

These visual reappearances are expressed cross-modally: we can hear Paul’s voice in units 20, 21 and 23. Since audiences have heard his voice earlier in the film, this is an explicit reappearance. However, there is a relation of difference with previous verbal cues as well, as he is now shouting and distressed, which translates into a very different voice quality. Moreover, the verbal cues referring to Paul are interrupted: after unit 21 ‘Come on now, three, four’ there is a brief interruption (unit 22) in the verbal-aural mode, then the count picks up again with ‘Five, six’, referring anaphorically to the first part of the enumeration, and continuing the aural-verbal link identifying Paul in his doctor’s coat, across the brief interruption. The interruption in the aural identity chain of Paul is due to the brief appearance of another character chain (his son Mark is watching closely from behind the operating room door). After this dialogue turn, there are no other manifestations of the aural mode that can be used to track Paul’s identity chain.

Finally, the cohesion and identity tracking in this sequence (units 16 to 27) is supported cross-modally by the sounds and music. Even though the sounds and music here have no relation of co-referentiality to Paul’s identity and are not strictly part of Paul’s identity chain (but part of the setting or non-diegetic), the high volume of the sound (beeping of a heart monitor) and music (still the same piano music) put the aural mode into a salient position in the narrative. Especially the increased tempo and rhythm (attributes of the comparative system) again mirror the tempo and rhythm of the kinesic actions of Paul. What is more, the beeping and piano music continue across all 12 units—even over shot 20 where Paul is on the beach and no longer in the operating room—tightening the cohesive link between both sub-chains (past and present) of Paul’s identity chain. What is remarkable is that the piano music in this sequence is largely non-diegetic (it is added as a sound effect during post-production to serve as a sound-bridge), but in unit 8 it is suddenly identified as diegetic, since Paul asks the nurse to play Bach in the operating room during surgery. This links the music more tightly to Dr Flanner as a kind of leitmotif.

Tracking the setting in the original clip

In the previous section we have focused on the detailed discussion of how the protagonist Dr Flanner can be tracked in a selected clip of *Nights in Rodanthe*. However, it became clear during the analysis that it is not always possible to separate the tracking of the character from that of other elements in the narrative, namely objects and settings, since they interact on several occasions. To paint a more complete picture of the scope of the cohesive links operating in this clip, we now move on to briefly discuss the setting identity chain. Parallel to the two sub-chains in the identity chain of Paul (present and past), there are two related sub-chains in the setting chain: the beach setting and the hospital setting.

The beach setting in this scene is re-presented as a specific setting in the first few units (units 1 to 5), by a combination of explicit and implicit visual reappearances through shots of a balcony, the beach, the ocean and the grey sky above; these are all identities that can be retrieved directly or indirectly through meronymy. The setting is rendered cross-modally, as
it is accompanied by the sound of waves in the background. The comparative system also contributes to the cohesion in this chain, as the qualitative attributes (colour, mise-en-scène, lighting) of the different reappearances resemble each other.

The cohesive strategies described above are repeated throughout the clip to support the identification and reappearance of the beach setting in units 7, 9, 15 and 20 (explicit and implicit visual reappearance though meronymy, supported by the sound of waves and the repetition of qualitative filmic comparison techniques such as lighting and colour). However, two particular units stand out: unit 14 contains a prelude (see description character chain) and in unit 20 the beach is visible but there is no sound of waves (the sounds of the hospital setting accompany the visuals of the beach).

The second sub-chain re-presented in this clip is the hospital where Dr Flanner works. The hospital has been presented before in the film and is part of the mental model of the audience by now. Two separate locations in the hospital reappear in this sequence: a hospital corridor (units 6 and 8), and an operating room (units 10–14; 16–19; 21–17). The setting chains consist of a series of implicit, visual reappearances, whose identities can be established indirectly through meronymy: both the corridor and the operating room can be inferred to be part of the larger already known identity of the hospital. In the same way, the setting identification is supported by close-ups of particular parts of objects and characters in the settings that are part of the overall hospital and operating room setting: the chart a nurse is carrying in unit 6, the patient’s face under green cloth in units 12, 13, 18, 22, Paul’s headlight in unit 14, the heart monitor visible in units 17 and 19, the green scrubs and face mask of Dr Flanner and the nurses, and the equipment Paul is using in units 12 and 13. We would like to highlight here that the identity chains of the participants are highly intertwined in film: secondary objects and characters (of which one could also analyze the individual identity chain), are here seen to be part of the identity chain of the setting, based on their salience and their meronymic/hyponymic relations.

The identities of the two settings are expressed cross-modally, as the visuals are supported by sound effects: the almost inaudible voices and mumbling of the operating room staff in unit 18 and the beeping of the heart monitor across units 16 to 24. Specifically, the operating room is also cued by a visual-verbal sign (anaphoric reference), namely diegetic on-screen text in unit 8, where ‘operating room 1’ is written on the electronic doors. The comparative system also contributes to the identity tracking and cohesion of the setting, more specifically through the bright colours and high-key lighting that characterizes all the reappearances of the hospital setting and that contrast with the low-key lighting and soft colours of the beach setting.

As was also the case with the character chain, the piano music plays a cohesive role in the setting chain. The piano music forms a sound-bridge between the alternating settings, suggesting a sense of unity between them. This is supported by the tempo and rhythm of the music that mirrors the tempo of the character’s kinesic action, and is also mirrored in the high tempo editing of this clip.

**Tracking the participants in the version with AD**

AD focuses on replacing the visual mode of the narration by a succinct verbal description and therefore relies on different items from the presenting, presuming and comparative systems to track the identities of the participants. But in contrast to ‘traditional’ verbal discourse, analyzed by Tseng (2013: 31–36) in terms of cohesive reference and identification and conceptualized by Halliday and Hasan (1976) and Martin (1992) in more depth, the
descriptions have to co-exist with the non-verbal and verbal-aural modes of the audiovisual product and are therefore timed to create links with the sound effects, music and dialogue. It is through this careful timing and the judicious selection of verbal cues that now replace the visual mode that it creates cross-modal aural cohesive links and manages to produce a simplified version of the original identity chains while remaining largely, but not entirely, synchronous with the chronology of the narration. In the next paragraphs we have a close look at the identity chains in the AD. We discuss the character and setting chain together, since we have argued in the previous section that they are highly fused, and that is no different for the AD version.

In units 1 and 2 the AD presents Paul as a specific (not generic) and ‘known’ person by using anaphoric references (his proper name ‘Paul’ and the personal pronouns ‘his’ and ‘he’). The AD adds a time indication, namely ‘Dawn’. This more explicitly links the present scene to the temporal development of the film than the film’s visual mode, which works through implication, i.e. here through the juxtaposition of the previous scene in the inn, which we are not discussing here, and the beach scene, which follow each other chronologically, and through the indication of time by way of its visual use of light and colour, among other things (as we have seen when analyzing the use of the comparative system in the previous section). The verbal-aural rendering of the AD makes the anaphorical link to the passing of time explicit and replaces the visually rendered morning light.

In units 1 and 2 the identity chain of the setting is tracked by the use of anaphoric reference through lexical cohesion, namely the repetition of the nouns ‘balcony’ and ‘beach’, part of the setting that has been presented/described earlier in the film. The chains in the AD are expressed cross-modally and create a direct cohesive link with the aural, non-verbal modes, i.e. the ocean waves, Paul’s footsteps on the beach and the sand and simultaneously—in an indirect verbal prelude—with the increasingly rhythmic sound of Paul’s breathing. It must be noted that the aural non-verbal cues of the identity chains are, indeed, accessible to the new target audience. However, the AD partially covers these sounds, so they are less redundant, since they are not available equally long.

What the AD does not render are the meronymic relationships (bridging) of the subsequent shots in the visual character identity chain or the visual details of the setting identity chain provided by the filmic images in units 3, 4 and 5—which ‘show’ more than linearly ordered words in an AD can ‘tell’. Overall there are fewer ‘reappearances’ of both Paul and the two main settings in the accessible version of the film, which works on the implied understanding that if the AD does not specify that characters leave or that a setting changes, they will still be assumed to be present (this is also the case for units 7, 9 and 20, for instance, where short reappearances to Paul on the beach are not explicitly described, as we will discuss later). Since there are fewer reappearances in the accessible version, the ‘comparative’ links associated with these reappearances are also less redundant and more succinct than in the original version. For instance, low-key lighting and colour (that express a relation of difference with the other setting in the clip) are subsumed under the description ‘dawn’. The AD does explicitly include the quality attribute of Paul’s jogging suit (see description unit 1). Mentioning the jogging suit provides a lexical link to the action Paul is performing and the setting in which he is performing it, the beach, thereby creating a mental model combining character-action-setting in the minds of the listeners through cross-modally functioning tracking. Another cohesive element of the comparative system highlighted in the analysis of the original clip and recreated in the AD is the rhythm and tempo of the music that mirrors the rhythm of the action. The careful timing of the AD (‘He begins to jog, then race along it’) is timed right before the up-tempo music starts and the fact that the AD leaves a small pause
for the sounds (for a brief moment the piano music, the quick footsteps and the breathing can be heard rhythmically interacting) potentially creates a comparable cohesive link across all chain elements as in the original clip.

In unit 6, the visual identity chain of Paul switches to the hospital setting. The AD describing unit 6 has anticipated this switch at the end of unit 2 where it already refers to ‘a hospital corridor’, identifying Paul in it with an anaphoric reference (personal pronoun ‘he’). Paul’s breathing, connected to the beach scene, is still briefly audible in unit 6 but is soon overtaken by the up-tempo piano music which, as we indicated above, contributes to the relation of similarity between Paul’s reappearances in different settings and creates a sound-bridge between the character chains of Paul in the two settings. In other words, the identity chain in the AD (‘He walks along a hospital corridor accompanied by a team of people’) makes use of the aural modes as well: the bridging music mixed with the gradually fainter sound of Paul’s breathing, in order to skip the meronymic visual shots of Paul on the beach in units 3, 4 and 5, and render the character and setting chains of unit 6, in the hospital corridor, instead. This is a necessary and common AD strategy, often referred to as ‘foreshadowing’, since in unit 6 dialogue kicks in leaving no room for description. A similar tactic is used for the intrusions of the visuals and sounds of the beach identity chains in units 7 and 9. Once the hospital identity chain becomes the dominant narration, the beach chain intrudes by association only through the rhythmic mix of breathing and music, whereby the salience of one or the other sometimes varies. Meanwhile, in unit 6 the dialogue confirms the AD, indicating that the character chain of Paul is now developing in a hospital setting and that the atmosphere in this setting is hectic, as the music and breathing suggest. This time the AD does not mention the quality attribute of Paul’s different outfit (the doctor’s coat) nor the high-key lighting and bright colours, reducing the visual redundancy of the cross-modal storytelling of the original clip.

More redundancy is lost as the AD of units 8 and 9 also skips a brief return to the hospital corridor, taking its target audience straight to the operating theatre where Paul is now preparing a patient for surgery. In other words, the identity chain in the AD counts fewer cues. However, the essence from the character and identity chains is rendered briefly in the AD (Paul is identified anaphorically by ‘he’ and the hospital setting is activated in the audience’s mind via meronymy by mentioning the noun ‘theatre’ and ‘surgery’). In addition, it is supported cross-modally by aural non-verbal cues as well as the dialogues. The up-tempo music creates a sound-bridge, but is here also identified in the dialogue as diegetic (‘Bach today, Bach!’) and the rhythmic breathing intrudes in the hospital soundtrack, in the way the shots to the beach intrude on the visual image in the original clip. Overall, the visual mode is pared to its essence in its verbal rendering and the aural modes are called to the rescue. The cohesive link between breathing and beach-jogging or the comparative link between the character chain on the beach and the character chain at the hospital becomes less explicit as it becomes less repetitive. What does remain is the generally hectic rhythm of both settings, which still provides the required link in support of the AD.

This implied connection is further enhanced by information the AD then selects for the description in units 11, 12, 13 in the operating room: the identity chain of the setting is maintained but reduced to the glaring lights of the operating room (‘bright overhead lights blaze’) in which the presence of Paul (‘him’) is also confirmed. The more subtle differences in lighting colour and dress between the beach scene and the operating scene expressed through the comparative system in the original clip are lost.

Nevertheless, the contrasting relation between the two settings is then restored more explicitly by the AD in units 14 and 15. It juxtaposes, in two subsequent sentences, the
character chain of ‘Paul’ racing along the beach, and the chain of Paul in the hospital, which implicitly reappears through lexical cohesion, i.e. by repeating the noun ‘the patient’, who had been introduced as part of the hospital setting in the AD of unit 9: ‘Paul races along the beach. The patient’s heartbeat flat lines’. In other words, the cohesive link between the two chains is created by implication and is supported cross-modally in the AD: the combined working of the AD and the non-verbal-aural mode, more specifically, its alternation with as well as its mixing of the rhythmic piano music (identified as operating music but also rhythmically connected to Paul’s stressed breathing), the sound of Paul’s breathing itself and the sound of waves. The AD at this point again provides a verbal prelude (or foreshadowing) to the telling sound that is heard a fraction later in unit 16: the beeping of the heart monitor. Subsequently, from units 16 to 23 the AD leaves the tracking of character and setting entirely up to the aural mode: the increasingly loud beeping sound of the instrument it has identified indirectly as a heart monitor (the audience needs to infer that the beeping comes from a heart monitor, based on the information that the patient’s heartbeat flat lines), the rhythmic piano music and the dialogues which confirm through the voices of medical personal (previously identified by the AD in unit 9 as ‘a team of people’) that the patient (first mentioned in unit 9) has passed away.

The AD then picks up Paul’s identity chain in unit 24, saying that he tries heart massage (‘As he pumps his patient’s chest . . .’), which is confirmed by the dialogue spoken by Paul himself. As the rhythm of the piano music also grows more hectic, the AD links the character chain of Paul to that of the man looking in through the OR’s glass door: ‘he looks over his shoulder and sees Marc watching anxiously’. The implicit link provided through juxtaposition in the visual mode, which implies that Paul sees the other character looking in through editing, is rendered more explicit in the verbal rendering of the AD which actually states that Paul (‘he’) ‘sees Mark’, who has been identified in the film (and the AD) earlier. All in all, some of the redundancy used in the original tracking of identity and setting chains is lost, while the tracking itself remains intact—supposing that the target audience can make good use of the aural clues provided to reconstruct the story and the mental models of the narrative they have constructed based on previous parts of the film.

Tracking the participants in the version with SDH

The challenges for SDH are quite different from those described above: all aural modes become inaccessible and must be rendered visually, including music, as demonstrated by Neves (2010). However, that is something SDH hardly ever seems to achieve today, as SDH often does little more than identify speakers or state that music can be heard (for more on SDH see Further Reading section). The scene under analysis is a case in point. It contains an abbreviated version of the verbal-aural mode in units 6, 8, 18, 20–21, 22 and 25, and only one reference to the specific sound produced by the heart monitor in unit 14.

The SDH banks on the fact that the target audience has access to the visual mode and hence to the visual character and setting identity chains as described under previous subsections and can use the subtitles to supply the information given through the dialogues, even if they can no longer access the aural non-verbal mode of sounds and music. For the identity chains on the beach this means: the sound of Paul’s footsteps, of the ocean waves, Paul’s breathing, and the rhythmic piano music, as well as the varying rhythms of breathing/piano and the way they alternate and/or are mixed. For the identity chains in the hospital this means: the same piano music and its intensifying rhythm in units 26–27, the beeping of the medical instruments and its growing intensity (unit 16), the muffled voices of the medical
staff surrounding Paul and the aforementioned sounds accompanying the cuts to the beach scene, linking the temporally distinct but thematically connected identity chains of character and setting.

All the same, the SDH does manage to salvage the basic participant identity chains of the film and is helped in doing so by the visual non-verbal and limited visual-verbal modes already present in the film. The SDH thereby continues to track the participant identity chains of character and setting/objects cross-modally to some extent. It renders the essence of all aural-verbal modes in writing and thereby provides cohesive links cross-modally with the visual non-verbal mode. Paul is not identified explicitly in the subtitles (by for instance mentioning his name) since he is on screen when he talks, but the other speakers at the hospital are identified as ‘man’ or ‘woman’ (see units 6 and 8), whereas the visuals convey that they are medical personnel. In unit 6 the subtitled dialogue sets the stage for the operation: in unit 8 it indicates that Bach will be played during the operation and the words ‘operating room 1’ on the hospital door name the specific setting explicitly through the film’s own visual-verbal mode. In unit 18 the subtitles summarize the dialogue stating that the patient has passed away and that a defibrillator is being used to revive him or her, and in units 21, 22 and 25 Paul’s subtitled counting indicates in a simultaneous cohesive link with the visual mode that he is performing heart massage.

One sound effect is explicitly subtitled: an explicit and direct presentation of the beeping sound of the heart monitor in the subtitles with unit 16 (direct identification of the source of the sound through a noun ‘EKG’ and mentioning of the fast tempo). This functions as a verbal-visual prelude for the dialogue subtitled in unit 18, granted that the beeping is only mentioned once, briefly, while the beeping sound actually continues until the end of the clip. Then, in unit 25 the visual-verbal mode provides an explicit cross-modal cohesive link between the verbal and non-verbal visual modes since the subtitle renders the continuation of Paul’s counting, whereas in the aural-verbal mode, the muffled voices are hardly discernible at this stage.

It is also important to remember that filmic sounds (see Remael 2012, referred to earlier in the paper) are produced by specific diegetic sources, the people as well as the settings and objects in the film. They have a visual as well as an aural rhythm and the target group of SDH can see them. While in this clip the aural rhythm is not rendered accessible (with the exception of unit 16), the accompanying visual rhythm reflected in the kinesic action and editing remains accessible. In the beach chain, visual rhythm is provided through the movement of Paul’s feet, the ocean waves and the meronymic close-ups of Paul breathing heavily. In the hospital chain, Paul and colleagues are marching along the corridor or moving hectically about the operating room. The different meronymic shots of the glaring operating room spotlights have their own spatial rhythm as they frame the figure of Paul. Moreover, the meronymic close-up of his fast-moving hands while operating, the shot of the heart monitor and its flat line, the way Paul turns his head in alarm, and the return to the visual rhythm of his run along the beach leave little to the imagination. What is more, the audience also has access to Paul’s visual expressions, the alarmed look on the face of the character behind the hospital door (units 25 and 27), and Paul’s awareness of him as he looks over his shoulder (unit 26). To what extent the effect of the visually accessible ‘sounds’ and other visual information can compensate for the lack of access to the aural non-verbal modes is hard to determine.

The main losses are the tonality of the voices and sounds (intonation, pitch, loudness), which contribute to the atmosphere of the scene and the hectic piano music, which as a structural element is linked to the setting of the hospital scene only by the SDH and no longer...
links the participant chains across time. Considering the way in which this loss is compensated, it is fair to hypothesize that the tracking of all participant chains remains intact, even if the added excitement that is the domain of sound and music accompanying the images is lost.

Concluding remarks

This chapter aimed to introduce the main concepts used for the analysis of multimodal or cross-modal cohesion in film today and to illustrate how they can be used to analyze AVTs. More concretely, we have demonstrated how filmic multimodal or cross-modal cohesion functions and investigated to what extent such cohesion is recreated in the accessible versions of a clip taken from a film with AD and SDH, while testing the usefulness of the said concepts.

The discussion of the theoretical framework underlined that research into multimodal cohesion is still in its infancy, and much conceptualization and theorizing remains to be done. Indeed, a truly systematic and comprehensive analytical model is not yet available for our research domain, whereas existing concepts and terminologies often overlap. Nevertheless, we have combined a number of these concepts in our analysis, in a bid to propose a first ‘inventory’ of the ways in which cohesion can be created in AD and SDH. We feel that Tseng’s model shows great potential for unravelling the intricate ways in which the different modes and sub-modes of audiovisual texts and their translations interact to create a cohesive audiovisual product.

The analysis of the original clip of Nights in Rodanthe highlighted the most salient elements that are at the disposal of viewers to track the identity of the most narratively relevant participants, a basic condition for cohesion. The subsequent discussion in sections showed that the identity tracking of participants remains intact in the accessible versions of this clip, even if the chains work differently, and sometimes rely on different modes or use the presenting-presuming-comparative system differently.

Whether these ‘shifts’ in the identity chains of the accessible versions of the clip with AD and SDH lead to a loss of cohesion and alter the reception of the narrative, cannot be determined by text analysis alone. Other methodologies, such as reception studies, should be used to corroborate text analytical findings. However, we have indicated that the ‘gaps’ or losses incurred in one mode, visual or aural, do not necessarily have to be problematic. The analysis of the original clip revealed how redundant the filmic medium is (the same information is conveyed more than once, through multiple modes), and one of these modes remains accessible to the new audiences of the accessible versions. The visually impaired audience extracts information from the soundtrack and music, the deaf and hard of hearing audience makes use of the visuals. This process of information retrieval works in combination with the process of mental model building, briefly mentioned in the opening section. Both target audiences are aware of the information that is not accessible to them and they actively apply compensation strategies, also relying on external sources such as their knowledge of the world. As the analysis has illustrated on several occasions, the loss of redundancy in the cross-modal cues leads to slimmer identity chains, a greater need for reconstruction through implication and hence, possibly, to a greater cognitive effort for our two specific audiences. On the other hand, less redundancy, succinct description and focused subtitling may also shield the blind and deaf audiences from information overload and help them grasp the most narratively salient information.

We must conclude by indicating that on the one hand, the potential of the Tseng model has not been fully exploited in our analysis and that there is room for more research, but also
that the model has its limits. Tseng underlines the importance of actions and action chains in film, which can be analyzed through the transitivitiy system (2013: 108–145), but adding more details to the analysis was beyond the scope of the present chapter. Conversely, she admits that tracking participants yields one type of information but that ‘[t]o provide a complex description . . . further analytical methods accounting for other dimensions still need to be developed, such as methods for describing filmic resources which manipulate emotions of different camera uses which affect and constrain information concerning characters’ (ibid.: 109). The same can be said for music and sounds. We have analyzed how the presence of music and sound can support the identification of characters on the textual level, but these aspects of film contribute to the production’s (connotative) meaning level as well. Moreover, they are not only employed to convey the characters’ emotions, but also manipulate the audience’s emotional involvement in the film. In order to chart the effect of, say, shifts or losses in textual redundancy, theoretical and text analytical approaches to the study of audiovisual texts should be complemented with experimental and cognitive studies taking on a cognitive and experimental view to analyzing the reception of audiovisual texts and their translations.

Summary
Multimodal productions, which make use of verbal, visual and aural modes to create a coherent message, need to fulfill basic textual requirements, one of which is that of textual cohesion, in this case multimodal cohesion. This chapter analyzes the challenges involved when such texts are made accessible for visually or aurally impaired audiences (e.g. through AD and SDH), since in such texts the original multimodal cohesion is altered. This chapter first defines and discusses the concept of multimodality, multimodal cohesion and cross-modal interaction. It discusses the analytical tools for multimodal analysis developed to date within social semiotics on the basis of seminal publications by van Leeuwen (2005), Royce (2007) and Tseng (2013). The chapter goes on to explore how these concepts and, in particular, the concept of ‘cohesive chains’ developed by Tseng, can be applied in the analysis of accessible multimodal texts. This is exemplified in a detailed multimodal transcription and analysis of a scene from the film Nights in Rodanthe. It illustrates and evaluates the above-mentioned analytical tools and inventorizes the ways in which multimodal cohesion is established in AD and SDH—thus paving the way for the development of new multimodal research methods within AVT.

Further reading
Kress, G. and T. van Leeuwen (1996). Reading Images: The Grammar of Visual Design. New York: Routledge | This is an innovative and seminal work exploring the grammar of images, based on Halliday’s social semiotic approach to meaning-making in language.
Multimodality and audiovisual translation

out in the framework of the European ADLAB project (www.adlabproject.eu) and provides a comprehensive overview of the challenges of AD and research in this field.


Related topics

6 Subtitling for deaf and hard of hearing audiences: moving forward
8 Audio description: evolving recommendations for usable, effective and enjoyable practices
13 Spoken discourse and conversational interaction in audiovisual translation
14 Psycholinguistics and perception in audiovisual translation
21 Multimodal corpora in audiovisual translation studies
22 Eye tracking in audiovisual translation research
23 Audiovisual translation and audience reception
31 Accessible filmmaking: translation and accessibility from production

References

Aline Remael and Nina Reviers


**Filmography**