Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films

Nina Reviers, Aline Remael

TricS, University of Antwerp, BELGIUM

ABSTRACT

This article aims to explore the complex issue of multimodal cohesion in audiovisual texts and their translations. It focuses on the analysis of Dutch multilingual audio-described films that combine audio description (AD) with audio subtitling (AST), a type of audiovisual translation (AVT) that raises many issues relevant to multimodal cohesion analysis in general. The article first discusses some of the standard multimodal textual features of translated audiovisual texts. It then looks at audiovisual texts that combine AD with AST, surveying some of the relevant literature on multimodality and the functioning of "multimodal cohesion" as opposed to linguistic cohesion. Finally the article surveys film clips from two Dutch productions that combine AD with AST, *Oorlogswinter* [Winter in Wartime] (Koolhoven 2008) and *Süskind* [Süskind] (van den Berg 2012), analysing how multimodal textual cohesion operates in these texts and pointing out the potential challenges for visually impaired users (VIPs). Generally speaking, it seems that AD and AST incorporate different types of explicit and implicit cohesive links that should allow a VIP audience to create a coherent audio text. However, further research will need to investigate whether these cues are sufficient for a VIP audience to understand the film and to what extent apparent shifts in textual cohesion may affect the reconstruction effort required.

KEYWORDS: audio description, audio subtitling, coherence, cohesion, multimodality.

1. Introduction: audio description with audio subtitling

As text types continue to proliferate under the influence of technology, audience diversification and globalisation, their semiotic composition and multimodal features are becoming increasingly complex. Granted that any text is always by definition multimodal (novels consist of words, the visual-verbal mode, and pictures of images, the non-verbal visual mode), films - the type of text that concerns us at present - typically contain multiple

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

50

aural and visual modes, making them quintessentially multimodal. The term "mode" is used here in the definition given by Kress and Van Leeuwen (1996), who also distinguish the concepts of "medium" (the material form of the text, for instance DVD versus newspaper) and "sign" (a specific instantiation of a resource in a given (con)text; for instance a red nose, which may signify a drunk or also a clown). The modes that, together, constitute a film are: the aural-verbal (film dialogues and lyrics); aural non-verbal (music and sound effects); visual verbal (various types of on-screen text) and visual non-verbal (images) (Delabastita 1989; Kress and Van Leeuwen 1996; Remael 2001; Zabalbeascoa 2008). Meaning is not merely created by the simultaneous presence of several of these modes in the multimodal product, but, more importantly, by their (simultaneous) interaction. To achieve this multimodal interaction, filmmakers create different types of interactive ties between two or more modes. Together, these ties serve as cues for the users to reconstruct a coherent end product, in this case a filmic story.

However, all these filmic modes can in their turn make use of different (semiotic) resources. For instance, the visual-nonverbal mode may resort to the use of gesture and body movement in order to create meaning or signify, whereas the aural-verbal mode may make use of different natural languages. In fact, today's multicultural societies produce an increasing number of multicultural and multilingual films, for a variety of reasons (Díaz Cintas 2011; O'Sullivan 2011).[1] Concretely, this means that the aural-verbal mode of films often consists of different languages, i.e. different systems of semiotic resources, reflecting, for instance, the different nationalities or cultural backgrounds of their characters. This tends to complicate the generation of meaning as a result of multimodal interaction and hence the reconstruction effort required on the part of the users, as will be clarified in the next paragraphs.

The complexity of the multimodal functioning of films increases further in translated versions, even in the case of what Jakobson termed "translation proper" or interlingual translation. In some forms of audiovisual translation (AVT), for example subtitling, a supplementary mode is added to the film (a visual-verbal mode or text on screen). In dubbing, the original aural-verbal mode, the film dialogue, is substituted by another, translated aural-verbal rendering. In both cases, the addition (subtitling) or substitution

---

[1] For a recent discussion of the challenges of multilingualism and its translation see Serban & Meylaerts (2014).

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

51

(dubbing) entails a change in the multimodal interaction of the film.

When the issue is expanded to include what Jakobson termed "intersemiotic translation", whereby one or several semiotic systems, or "modes" in our terminology, are translated into another, the plot thickens considerably. This happens in audio description (AD), which makes audiovisual products accessible for the visually impaired by rendering images (visual non-verbal) and unclear sounds (aural non-verbal) in an aural-verbal narration, a form of substitution that is added to the sound track and interacts with the dialogues and sounds of the original text.

Due to the current globalised flow of information and entertainment this also means that the AD may have to be combined with other forms of AVT, closer to "translation proper" (cf. above), in order to function in different markets. In the case of a dubbed film, AD is combined with the dubbed dialogue. In precise terms this means that an American film dubbed into Spanish for example will be supplied with a Spanish AD that can interact with the dubbed Spanish dialogues to create a coherent Spanish audio-described film. However, in the case of subtitled films, the AD, an aural-verbal mode, cannot interact with the subtitles, a verbal-visual mode, since the core users of AD are blind or visually impaired. This also holds true for multilingual films, in which subtitling is used to make scenes in a language other than the main language of the film, accessible to viewers who do not know the language in question. For visually impaired users (VIPs), this means that the written subtitles, a visual-verbal mode, must be given an aural-verbal rendering through what is known as audio-subtitling (AST).

AST can therefore be defined as the aurally rendered and recorded version of the subtitles with a film. This spoken version of the subtitles is mixed with the original sound track. AST is usually read, sometimes acted out, by one or more voice actors. Sometimes it is produced by text-to-speech software. The subtitle text is often delivered almost literally, but it can be rewritten to varying degrees, and in addition, the recording method also varies. Usually, AST is recorded as a form of voice-over, which means that the original dialogues can be heard briefly before the translation starts. Sometimes it is recorded in a semi-dubbed form, which means that the original dialogues are substituted by a form of dubbing that is not necessarily

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

52

entirely lip-sync, that is, synchronous with the lip movement of the speaker.[2]

Of interest in this context is that each and every AVT variant implicates different requirements for the multimodal functioning or signifying of the text and for the cues placed at the disposal of the user to re-create a coherent story. More specifically, the combination of AD with AST has a substantial impact on the functioning of the text as a coherent whole since the cohesive devices that serve as cues to create coherence between the image and the dialogues in the original film need to be translated into verbal-verbal cohesive links between the AD and the AST in the audio described version in order to maintain the same semantic unity.[3]

The aim of this study is therefore to discuss the devices used to ensure multimodal cohesion in multilingual films with AD and AST. Our central research question is: What type of cohesive devices are used (in AD) to (re)create cohesion with the translated dialogues or AST? This question is explored through the analysis of two sample clips from multilingual Dutch films combining AD with AST. In addition, we will point out what appear to be ensuing shifts in the thus (re)created cohesion of the material analysed, suggesting avenues for further research into the effectiveness of AD-AST combinations.

Before moving on to the discussion of the clips under analysis and a series of examples in sections 3 and 4, section 2 examines some of the methodological and terminological aspects involved in analysing multimodal cohesion in AD with AST.

2. Defining and analysing multimodal cohesion in AD

The multimodal approach to text analysis is an emerging paradigm within Linguistics and Translation Studies. However, there are still strong methodological challenges regarding how to define and analyse the construction of multimodal meaning and how to tackle its

---

[2] For more information on AD-AST combinations see also Braun and Orero (2010); Remael (2012) and www.adlabproject.eu.

[3] In this context the concepts of cohesion and coherence will be defined as follows. Cohesion is a textual property that is achieved by the whole of all grammatical, lexical and other textual links or cohesive devices inserted into the text by the author in order to hold the text together. These devices allow for coherence to be created. Coherence is understood to be the result of the mental process of making sense of a text, multimodal or other, by the users, on the basis of their background knowledge and disposition as well as the communicative context and the cohesive devices present in the text (Halliday & Hassan 1976, Braun 2011, Tincheva 2012).

translation (see for example Chaume 2002; Gambier 2013). A driving force behind multimodality is the Social Semiotics tradition sparked by Kress and Van Leeuwen (1996) in the late 1990s, which was rooted in the Systemic Functional Linguistics (SFL) approach of Halliday and Hassan (1976). The former combines SFL concepts with insights from Film Studies (for example, Bordwell and Thompson 2010). Since then, several linguists and translation scholars have shown that some of the concepts and analytical tools of SFL and Social Semiotics indeed have potential for explaining the construction of multimodal meaning and cohesion (see for example Baldry 2005; Baumgarten 2008; Braun 2011; Norris 2004; Royce 2007; Sindoni 2011). The present study aims to test and build on this emerging tradition by conducting an analysis of cohesion in two film clips with AD and AST, starting from a general, pragmatic definition of multimodal cohesion as any instance in which an implicit or explicit "sense-relation" exists between two or more signs of a different or same mode in a given text that helps the viewer to create a coherent textual semantic unit.

In order to carry out textual analyses, however, the definition needs more precise elaboration. When consulting the as-yet-limited body of work on Multimodality, the authors of this article distinguished (at least) two inter-related types of multimodal cohesion between image and dialogue in films, even though many different terms and definitions are used to refer to these cohesive ties in research literature. Firstly, there is the concept of cohesion through reference: two signs in a text refer to each other in such a way that the interpretation of one cannot be successfully realized by the viewer without the presence of the other in the text or in its immediate context. Most often this concept refers to visual-verbal deictic references such as: "What is that?" where "that" refers to an element (the referent) present in the image. Baumgarten (2008:14), for example, analyzes the cohesive links between images and dialogues in a series of films and identified traditional verbal cohesive devices such as personal pronouns, demonstrative and possessive pronouns and determiners, definite and indefinite articles, adverbs of time or proper names. However, she also underlines the importance of more implicit sense relations between modes due to their simple co-occurrence, and that brings us to a second type of multimodal cohesion between visual and verbal signs, which we will call complementarity (following Royce 2007).

Signs in a given multimodal text often occur simultaneously or in each other's immediate

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

54

textual environment and create a sense of semantic unity. Royce (2007:103) defines complementarity in multimodal texts as instances in which visual and verbal modes co-occur to complement each other semantically and produce a single textual phenomenon. He applies Halliday's concepts of lexical cohesion to identify the cohesive devices that are used to create such complementarity. The most important of these are repetition, synonymy, antonymy, hyponymy, meronymy and collocation. The idea of complementarity between modes also (re)appears in different works on AVT, each time in slightly different guises: from Marleau's (1982) functions of subtitles (*ancrage*, *relais* and *redundance*), and Zabalbeascoa's (2008) interactive relations between modes (complementarity, redundancy, contradiction, incoherence), to Gambier's (2013:48) "relationships between sound, picture and verbal content" (redundancy, complementarity, autonomy, contradiction, distance, criticism, help).

When consulting the few existing articles specifically addressing coherence and cohesion in AD, a subdivision emerges comparable to the one discussed above between explicit (textual) links and more implicit relations. Yos (2005), for example, analyses the cohesion between AD and dialogue in the German *Tatort* [Crime Scene] episode *Laura, Mein Engel* [Laura, my angel] (Runze 1994) and concludes that the cohesion between image and dialogue is mainly expressed through the explicit identification in the AD of speakers and objects. The speaker identifications can be made directly through, for example, deictic references (e.g. personal pronouns and proper names) and objects can be identified by describing them through lexical items. Braun (2011), for instance, gives an example from the AD for *The Girl with the Pearl Earring* (Webber 2003) in which the dialogue contains the pronoun "it", referring to what is described in the AD as "a painting".

However, Yos also identifies several indirect cohesive links for character identification, based on inference by the audience: inference from the dialogue, sound or from the AD. For example, Yos emphasises that AD dialogue cohesion is strongly determined by the description of actions in the AD, since dialogues often comment on or relate to what is happening on screen. Simply describing what is happening or where characters are looking therefore usually provides the audience with enough information to infer who is speaking. Braun (2007) makes the same observation in her analysis of *The Girl with the Pearl Earring*, stressing that many visual signs are produced in response or in reaction to verbal utterances

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

55

or vice versa. To maintain cohesion, describing the ongoing action is often sufficient for the audience to infer the meaning of a specific, related verbal response.

Finally, Braun and Orero's (2010) survey of ten English-language films combining AD with AST reveals that in these films the ADs contain more explicit speaker identifications than "regular" ADs which work in interaction with the original dialogues. In addition, and depending on how the AST was recorded from a technical perspective, they identify broadly the same cohesive devices for speaker identification as Yos: direct identification, as in the case of direct speech, or indirect identification by describing what characters do or what they are looking at.

Nevertheless, the above discussion of multimodal construction of meaning and cohesion in film and AD still does not provide us with a clear or comprehensive framework for analysing multimodal cohesion in AD. Further conceptualisation is still required to create a comprehensive translation-oriented approach to multimodal text analysis, which is beyond the scope of this article. However, the discussion highlights a number of useful concepts that can serve as the basis for our initial analysis of AD-AST cohesion: reference, lexical cohesion and "complementarity", an umbrella term that we will use to group all the multimodal cohesive devices described in the literature which provide indirect semantic links between signs of different modes. Having said that, we are fully aware that some degree of interpretation is unavoidable in assigning linguistic concepts, such as (lexical) cohesion and reference, to a given textual feature.

3. Data collection and methodological approach

In section 4 we will examine the multimodal cohesion of two clips from multilingual Dutch film productions that make use of AD with AST: *Süskind* [Süskind] (Van den Berg 2012) and *Oorlogswinter* ([Winter in Wartime] Koolhoven 2008). Both films take place in the Netherlands during the Second World War. In *Süskind*, the protagonist, Walter Süskind, a member of the Jewish Council of Amsterdam, finds a way to safeguard children against deportation to the death camps. In order to do this, he falsifies transportation figures and becomes friends with a German SS officer, Ferdinand Aus der Fünten. However, the latter is told about the deceit at some point and raids the theatre and kindergarten that are used as a

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

56

transit venue by the Council. One of the other characters mentioned in the examples below, Bauer, is another SS officer. *Oorlogswinter* takes place in a small town in the Netherlands, where protagonist Michiel van Beusekom, a 14-year-old boy, lives with his father Johan, the mayor of the town, his mother Lia and older sister, Erica. In the scene we are about to analyse, Johan has been arrested by the SS and Lia and Michiel are rushing across the village to the town hall where he is being held in the hope of seeing him and delivering some food for him.

The choice of the films is based on several criteria. Firstly, there are only very few films available with AD and AST in Dutch, the language that is the focus of our analysis, and only *Oorlogswinter* and *Süskind* use AST throughout the whole film. The other four productions that have foreign language dialogues with AST - *De Zaak Alzheimer* [The Alzheimer Case] (Van Looy 2003*), Dossier K* [The K files] (Verheyen 2009), *Tirza* [Tirza] (Van den Berg 2010) and a few episodes of the TV series *Wolven* [Wolves] (Desmyter 2012-2013) - use AST only occasionally. Secondly, *Oorlogswinter* and *Süskind* are of the same genre, both are World War II films. Thirdly, in *Süskind* the producers opted for semi-dubbing; the written subtitles of the scenes in which German and Dutch are mixed have been rewritten and expanded to resemble the original film dialogues, before being voiced. In *Oorlogswinter* by contrast, the AST was recorded in voice-over mode and the subtitles were only slightly adapted. Consequently, together, the films demonstrate the two methods currently being practised for recording AST in the Low Countries. In both films, the subtitles are "acted out" by voice-actors, which is also common practice for fiction film with AD/AST in Flanders and the Netherlands. Fourthly, the two clips selected for our analysis, are particularly representative of the two films in terms of the language rendered by the AST (German) and the characters that feature in the scenes (the protagonists).[4] Finally, all the filmic modes that can theoretically play a role in the signifying process are evident in the scenes: dialogue, image, music and soundtrack.

---

[4] Both excerpts were taken from the (region 2) DVD and converted to .mp4 format. The timecodes for the clip from Oorlogswinter (Koolhoven 2008) are: 00:54:51:16 to 00:56:08:18. The timecodes for Süskind (Van den Berg 2012) are: 01:18:11:00 to 01:19:09:11.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

The analysis of the clips was conducted in two steps. Firstly, we made a basic multimodal transcription of the selected clips in an excel file. Transcribed categories include timecode, lower-level actions (see Norris 2004), visual film techniques, original dialogue, AD, AST, sound effects and music. This initial step allowed us to identify all the possible instances of AD-AST and AD-sound interaction in the two clips. Next, a few especially representative excerpts from this transcription that illustrate each of the different types of cohesive links discussed in section 2 (reference, lexical cohesion and complementarity) were selected in order to carry out an in-depth analysis of the functioning of these types of cohesion in our particular text type. The said excerpts constitute the dataset discussed in the next section.

4. Discussion and analysis of AD-AST cohesion

In the clips under analysis we identified several instances of meaningful AD-AST interaction during the transcription phase. It also appeared that all the relations identified in the clips could be allocated to the three categories mentioned in the literature: reference, lexical cohesion and complementarity. Both clips feature all three categories of cohesion, even though some sub-categories featured more often in one clip than in the other. This will be made clear in the next paragraphs, where we analyse representative examples from the clips for each cohesion category in order to illustrate and discuss the way they function in detail.

4.1 Reference

A number of cohesive links between AD and AST established through the use of reference are evident in both clips. In all the instances of reference in the clips under analysis, the cohesive relation is established through the interaction between a referring item in the AST (personal pronoun, proper name, possessive, etc.) and its referent, rendered verbally in the AD in lieu of the original visual information, that is, a lexical item (such as a noun) or another form of referent/antecedent (proper name, adverb, etc.). The references that occur in the clips are both anaphoric and cataphoric, and the referents are almost always characters rather than objects. There are instances of referring items with a clear/explicit referent and some with more implicit or even missing referents. A few representative examples of each of these types of reference from our *Süskind* and *Oorlogswinter* clips are discussed below. For all the examples we have included a transcription of the surrounding AD and AST in order to provide the necessary context.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

58

4.1.1 Explicit referents

The scene transcribed in Table 1 takes place in the theatre that serves as a transit centre for Jews about to be deported (see section 3). The Jews are invisible at first because they are hiding between the rows of seats.

Table 1: Excerpt from *Süskind* (part 1).

Dialogue/AST	Audio Description
	Aus der Fünten stapt verder en loopt het podium op. [**Aus der Fünten** proceeds and walks onto the stage]
AUS DER FÜNTEN: Süskind? Süskind? Waar zijn **mijn** Joden? [AUS DER FÜNTEN: Süskind? Süskind? Where are **my** Jews?]	
	Bauer ziet <u>iemand</u>. [Bauer sees <u>someone</u>]
BAUER: Hier. [BAUER: Here]	
AUS DER FÜNTEN: Hé, <u>jij</u> daar, opstaan. [AUS DER FÜNTEN: Hey, <u>you</u> over there, get up]	
	<u>Een vrouw</u> staat op. [<u>A woman</u> stands up]
AUS DER FÜNTEN: Waar is Süskind? (STILTE). **Ik** heb je wat gevraagd. Breng *ze* naar buiten, Bauer, *allemaal*. Eruit. [AUS DER FÜNTEN: Where is Süskind? (SILENCE). **I** asked you something. Take *them* outside, Bauer, *all of them*. Out.]	
	Meer mensen volgen. [*More people* follow]

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

BAUER: Naar buiten zei ik. Zijn *jullie* doof? Naar buiten. [BAUER: Outside, I said. Are *you* deaf? Go outside.]	

This excerpt from the AD for *Süskind* contains several instances of references:

Example 1: The AST contains referring items denoting the main character Aus der Fünten, such as possessives and personal pronouns (indicated in bold type in Table 1). The AD identifies their referent by introducing the character by his proper name (also in bold type).

Example 2: The AD then introduces a new character with the pronoun "iemand" [someone], the AST contains referring items denoting this referent, i.e. the personal pronouns "jij" and "je" [you] (underlined in Table 1), creating cohesion without giving away who the "someone" is. The person is later, cataphorically, identified by the AD as "een vrouw" [a woman], once the character's action identifies who she is (underlined words in Table 1).

Example 3: The AST contains referring items in the form of personal pronouns "ze"[them], "jullie"[you, pl.] and an adverb "allemaal" [all of them], and the AD cataphorically identifies their visual referent with a general noun "meer mensen" [more people] (in italics in Table 1).

In the following excerpt from *Oorlogswinter* (Table 2), Lia and her son Michiel arrive at the City Hall (see section 1).

Table 2: Excerpt from *Oorlogswinter*

Dialogue/AST	Audio Description
	Lia klopt op de deur. **De soldaat** die open doet, knikt hen beleefd toe. [**Lia** knocks on the door. **The soldier** who opens (it), nods politely to them]

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

LIA: Goede morgen. **Ik** zou graag **mijn** echtgenoot willen spreken. Van Beusekom. [LIA: Good morning. **I** would like to speak to **my** husband. Van Beusekom.] SOLDAAT: Het spijt **me**, **ik** kan **u** niet toelaten. [SOLDIER: **I**'m sorry, **I** cannot let **you** in.] LIA: Maar **ik** heb kleding en eten bij **me**. [LIA: But **I** have clothes and food with **me**.] SOLDAAT: Dank **u** wel, **ik** zal ervoor zorgen… [SOLDIER: Thank **you**, **I** will see to it…	
LIA: Nee, dat wil **ik** hem zelf geven. Alstublieft. [LIA: No, **I** want to give them to him myself. Please.]	
	Ineens duwt Lia de soldaat aan de kant en schiet naar binnen. [Suddenly Lia pushes the soldier aside and forces her way into the building.]
SOLDAAT: Het spijt **me**, **ik** mag **u** echt niet binnenlaten. [SOLDIER: **I**'m sorry, **I** really can't let **you** in.]	

The above excerpt from *Oorlogswinter* contains many comparable instances of the use of references ensuring cohesion between AD and AST.

Example 4: The conversation between Lia and the soldier who opens the front door of the City Hall is rife with personal pronouns and possessives referring to each of them (indicated in bold type in Table 2 and in the corresponding forms in the back translation). In each case, the AD identifies the referent anaphorically, by a proper name (Lia) in one case, and by a noun (the soldier) in the other.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

61

4.1.2 "Implicit" referents

The instances where the referent denoted by the AST is not as explicitly evident in the AD as in the previous examples is especially interesting. In the clips under investigation, we identified the following instances of implicit or even absent referents, grouped into three types as indicated below.

Implicit reference type 1

Sometimes a referring item in the dialogues and AST denotes a character who is not physically present in the filmic world at the time of speaking, but was introduced into the narrative earlier and is therefore supposedly known to the audience or part of the mental narrative model that audiences construct when interpreting a narrative. As Emmott and Alexander (2009:412) write: "Linguists, psychologists and narrative scholars employ schemata theory to account for the interpretation of a text where the discourse itself does not provide all the information necessary for the discourse to be processed." Johnson-Laird (1983:141) uses the term "mental model" to refer to the functioning of such schemata, arguing that we all make sense of the world around us by constructing mental models of reality. The same occurs when readers or viewers make sense of fictional worlds: They construct mental models of the fictional world that is being portrayed before their eyes using both bottom-up information from the text and top-down information, tapping into their personal world knowledge (Vercauteren and Remael 2014). In other words, in the current example, too, the referent can be "inferred". He or she has appeared in previous scenes and has been mentioned explicitly by the AD at the time but there is a considerable time lapse between the present occurrence of the referring item in the AST and the referent in the AD. We give an example of implicit cohesion type 1 from each clip below.

Example 5 (from *Süskind*): In the opening sequence of the clip, Aus Der Fünten is calling for Walter Süskind by name: *"Süskind? Süskind? Waar zijn mijn joden?"* [Süskind? Süskind? Where are my Jews?] (see transcription Table 1). However, Walter Süskind is not present at that moment, which is clear from Aus Der Fünten's behaviour. He is therefore not mentioned in the AD. Süskind is, however, part of the mental model of the audience. He has been in the setting in which the present scene is taking place and has been identified as the director of the transit venue for the Jews who are to be deported.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

62

Example 6 (from *Oorlogswinter):* In her exchange with the soldier who opens the door of the City Hall, Lia refers to her husband using nouns, his proper name and personal pronouns – "my husband", "Van Beusekom", "he", "him" etc. (see transcription Table 2). Johan is not physically present in the scene at the door, but he has already been introduced in the narrative as Lia's husband. The audience therefore knows the character and is aware that he is being held by the Germans in the City Hall, which Lia is trying to enter.

Implicit reference type 2

Sometimes the referent of a referring item in the AST is simply "unidentified" in the AD. In other words, the referent is not explicitly mentioned in any way or form, but recipients can (theoretically) infer the referent from the context. This type of implicit reference only occurs once in the clip from *Süskind*, as exemplified below.

Example 7: The locative demonstrative adverbs (here/there) used in the AST in the above excerpt from *Süskind* (BAUER: **"Hier"**. AUS DER FÜNTEN: "Hé, jij **daar**, opstaan" [BAUER: **Here**. AUS DER FÜNTEN: Hey, you **there**, get up.]; see transcription in Table 1) do not correspond to a location or a referent described in the AD. Strictly speaking, the referent is therefore missing, which results in an instance of incoherence. However, the AD does describe the action associated with the utterance, namely "Bauer ziet iemand" [Bauer sees someone], without identifying who the "someone" is. As a result, the user can infer that "hier" [here] and "daar" [there] imply two locations that are close by and visible to the characters who are speaking, and that the adverbs refer to these places because the other, unseen characters are hiding there. However, this may require more effort on the part of a non-sighted audience.

Implicit reference type 3

In a few cases in the clips under investigation, the AD referent denoted by a referring item in the AST is simply unidentifiable; again, this type of implicit reference only occurs in the excerpt from *Süskind*. In the excerpt transcribed below (table 3), Süskind is in his office, talking on the phone.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

63

Table 3: Excerpt from *Süskind* (part 2).

Dialogue	Audio Description
	In Walters kantoor [In Walter's office]
WALTER: Dat zeg ik. Luister, M. Cohen, ik verzeker u, **het** is mij ook een raadsel. Ja, één moment, één moment. [WALTER: That's what I am saying. Listen, Mr. Cohen, I assure you, **it** is a mystery to me too. Yes, hold on, hold on.]	
	Hij legt de hoorn neer en loopt naar het raam. Duitse soldaten drijven de mensen de straat op. [He puts down the receiver and runs to the window. German soldiers are driving people into the street.]
WALTER: Oh God. [Oh God.]	

Example 8: In this sequence from the *Süskind* clip, Walter is on the phone and refers to something using the definite article "het" [it] in his conversation with an invisible interlocutor (see Table 3: bold type). There are no cues in the AD or context to help identify the referent because Walter's interlocutor cannot be heard or seen, and the VIP audience cannot know what Walter and this invisible person are talking about, but then again, this is no different for the sighted audience. Both the sighted and the blind audience must assume what is happening from the context; this is a telephone conversation with an invisible and inaudible speaker on an unidentified topic. That Walter is on the phone is loosely communicated by the AD, retrospectively ("hij legt de hoorn neer" [he puts down the receiver]) and can possibly be derived from the one-sided form the conversation takes.

4.2 Lexical cohesion

We have already discussed some lexical items such as nouns and adverbs when dealing with reference in the previous section. However, our clips also contain other instances of lexical cohesion such as reiteration and, to a lesser extent, collocation. As in the previous section, we looked at instances where a lexical item from the AST is picked up by the AD. Examples of each type of lexical cohesion are discussed below.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

64

4.2.1 Reiteration

Lexical items from the AST are quite frequently reiterated in the AD in order to maintain cohesion, even though they take different forms and have different degrees of explicitness. The excerpt from *Süskind* contains most of this type of cohesive devices, as our examples show, but reiteration is used in both films.

In some cases simple repetition of the same lexical item occurs.

Example 9 (Süskind): In this example the lexical item "opstaan" from the AST is repeated in the AD. After Aus der Fünten's turn: "Hé, jij daar, **opstaan"** [Hey, you there, **stand up**]. The AD says: "Een vrouw **staat op"** [A woman **stands up**]. (See *Süskind* Table 1 for context).

Example 10 (Süskind): In much the same way, the lexical item "two" from the AST is repeated in the AD, as is indicated in bold type in the transcription of the third part of the *Süskind* clip below (Table 4).

Table 4: Excerpt from *Süskind* (part 3).

Dialogue	Audio Description
GENERAAL: Bauer! BAUER: Jawel [Yes] GENERAAL: Neem **twee** *mannen* en maak de overkant leeg. [GENERAL: Take **two** *men* and empty the other side.] BAUER: Tot uw orders. Jij, meekomen, jij ook. [BAUER: At your command. You, come along, you.]	
	Bauer loopt met **twee** *soldaten* naar de crèche. [Bauer walks over to the kindergarten with **two** *soldiers*.]

	Walter haast zich de trap af.
	[Walter rushes down the stairs]
	Buiten staan de *mensen* naast elkaar met hun
	handen omhoog geheven.
	[Outside people are standing in lines with their
	hands raised.]
AUS DER FÜNTEN: Waar is de rest van mijn 7.000 *Joden*? [Where are the rest of my 7,000 Jews?]	

In a number of instances in our clips, the AD and AST contain lexical items that are each other's hyponym or hypernym. Sometimes the AD will contain the hyponym (the more specific term), sometimes the AD will contain the hypernym (the more general term) to refer to the same referent. The purpose of this combination of more and less specific terms appears to be the creation of a "new" aural cohesive text that functions independently, relying on AD – AST interaction.

Example 11 (Süskind): Aus der Fünten is looking for the missing Jews and says "Waar zijn mijn **Joden**?" [Where are my Jews?], when the AD indicates, "Bauer ziet **iemand**" [Bauer sees **someone**]. This "someone" is subsequently identified in the AD as "a woman" in "Een vrouw staat op" [A woman stands up] (see Table 1). This woman is part of the larger group of "Jews" that Aus der Fünten has referred to.

Example 12 (Süskind): The AD uses the more specific term "soldaten" [soldiers] in "Bauer loopt met twee **soldaten** naar de crèche. [Bauer walks over to the kindergarten with two **soldiers**.]" rather than "mannen" [men], which is the noun used in the dialogue by the general in "Neem twee **mannen** en maak de overkant leeg" [Take two **men** and empty the other side.] (see italics in Table 4).

Example 13 (Suskind): In our next example, (also from Table 4) the dialogue/AST refers to the people present on screen as "**Joden**" [**Jews**], when the general shouts, "Waar is de rest van mijn 7.000 **Joden**" [Where are the rest of my 7,000 **Jews**?]. The AD, in contrast, describes the groups of Jews more generally as "**mensen**" [**people**] in "Buiten staan de **mensen** naast elkaar met hun handen omhoog geheven." [Outside the **people** are standing in

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

66

lines with their hands raised.].

Example 14 (Oorlogswinter): In the clip from *Oorlogswinter*, Lia says that she has brought "kleding en eten" [food and clothes] for her husband (see Table 2), pointing to a basket that she is carrying. The basket is not mentioned in the AD, most likely due to time constraints, but it has been identified earlier in the scene as "een mand vol **spullen**" [a basket full of **things**], the "spullen" being a general noun for what is later identified as "**kleding en eten**" [**food and clothes**].

A final example of lexical cohesion that merits special mention comes from the *Süskind* clip.

Example 15 (Süskind): In this example the words "de overkant" [the other side] and "de crèche" are used as synonyms to refer to the same referent. However, the noun used in the AST does not fully identify what the speaker is referring to except in spatial terms: "de overkant" [the other side], which only the viewers can see. The AD fills in the referent with a more specific lexical item "de crèche" [the kindergarten] thereby rendering cohesion with previous scenes more explicit (see Table 4).

It has become clear from the examples that reference and lexical cohesion are concepts that work closely together to create textual cohesion. Together, they create cohesive threads across the AD and AST texts. Examples 2 and 11 from *Süskind* offer a fitting illustration of such a thread with items that refer to the same referent in the narrative, namely the Jewish woman. In addition, the cohesive thread involving the woman forms part of an even longer thread that takes in all the Jews in the clip, a group to which she, in fact, belongs. In other words, the first thread: - (iemand [AD/someone] – jij [AST/you]– een vrouw [AD/a woman]– je [AST/you]) - is followed by a second thread that is linked to the first - (ze [AST/them]– allemaal[AST/all] –de mensen [AD/people] – jullie [AST/you, pl.]– mijn Joden [AST/my Jews]) (see Tables 1 and 4 for context).

4.2.2 Collocation

Another type of lexical cohesion is collocation, a category that is more difficult to pinpoint because it is a much vaguer concept. Halliday and Hassan define it as two lexical items that are connected to each other in some word-meaning relation (1976:285). Such collocations of

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

67

lexical items linking AD and AST occurred only rarely in the clips under analysis, but this does not mean that they might not be more numerous in other films combining AD and AST. We identified one instance of collocation in *Süskind.*

Example 16 (Süskind): The collocation consists of the logical, lexical combination of an adverb and a verb. The adverb "naar buiten/eruit" [outside/out] (in the AST), which is expressed as "Outside, I said. Are you deaf? Go outside", can logically be expected to co-occur with the verb "volgen" [follow] (in the AD), because it verbalizes the visual movement of the action (see Table 1).

5. Cohesion via "complementarity"

We indicated above (sections 2 and 4.3) that the concepts of referential and lexical cohesion are probably insufficient to analyse the way in which cohesion in AD does or does not work. Indeed, several scholars researching the fields of AVT and multimodality have identified complementary relations between signs of distinct modes that are of a different nature and cannot be captured by the linguistic (Hallidayan) concepts of cohesion alone (see section 2).

The clips from *Süskind* and *Oorlogswinter* both contain a number of such instances and illustrate that the redundant nature of filmic narration is essentially the element that lies at the heart of this extra-linguistic cohesive mechanism. Film makes the most of images, sounds and dialogue to bring its story across, employing what we will term "multimodal redundancy", created by the dialogic interaction between the different filmic modes. In the representative examples of cohesion through complementarity discussed below, the dialogues are a reaction or response to signs conveyed by other modes, visual or aural. In other words, they represent instances discussed above in the introduction (see section 1), where we refer to Yos (2005) and Braun (2007): both scholars point out that film dialogues often comment on or relate to actions or events happening on screen, adding that it is therefore important for the AD to describe these actions in order to ensure that textual cohesion across modes is maintained.

Such interactions are not always signalled explicitly by linguistic elements in the original film dialogues. Instead, they rely on inference. This also occurs in the AD/AST version. As

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

68

the examples below demonstrate, the combination of visuals or sounds with dialogue provides sufficient contextual information for the audience to infer coherence, without the need for explicit cohesive devices. When AD is combined with AST, the complementarity between the AD and the dialogues (itself replacing the complementarity between visuals/sounds and dialogues) must be replaced by a comparable complementarity between the AD and the corresponding information in the AST. We refer to Tables 1, 2, 3 and 4 for additional context with the examples below.

Example 18 (Süskind): Aus der Fünten orders a woman who had been hiding between two rows of seats in the theatre to stand up and then asks her where Süskind is. When she does not react or respond, he proceeds to say: "Ik heb je wat gevraagd." [I have asked you a question]. There is a clear cohesive sense-relation between the woman's obvious (in)action and Aus den Fünten's lines of dialogue in the original film, which is not picked up by the AD. On the one hand, Aus der Fünten's retort is clearly provoked by the woman's silence, which is accessible to the VIP audiences as much as it is to the sighted audience. On the other hand, there is a form of multimodal redundancy in the film that is lost in the AD/AST version. The sighted audience can hear the silence but can also see that the woman does not react, in a revealing shot of the woman's blank face. This redundancy is lost in the AD.

Example 19 (Süskind): When Süskind is on the phone, speaking to an unknown interlocutor, he suddenly says "één moment [hold on]" and puts down the receiver in reaction to sounds he hears in the street where a number of Jews are being lined up by German soldiers. Again a logical sense-relation pertains between dialogue and action (Süskind says, "hold on" because he needs to stand up to walk over to the window). Sighted viewers can hear the sounds in the street that prompt Walter to stand up from his chair but they can also see him glance at the window from his seat behind the desk and subsequently stand up. This dialogic interaction between dialogue, sounds and action is recreated in the AD. Walter's action of standing up is described; the VIP audience can also hear the dialogues (AST) as well as the noises that Walter and the sighted audience hear. Nevertheless, the AD audience again misses out on some of the multimodal redundancy of the film as they are not told that Süskind first glances at the window. In short, they have to rely on aural input only, some of which is not easily identifiable (i.e. the noise in the street) and it is not explained in the AD/AST version of the

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

69

film.

Example 20: Another example in which redundancy possibly makes it "easier" for the sighted audience to infer the meaning of a sequence is Walter's telephone conversation itself (see Table 3). The VIP audience can only infer in retrospect, when the AD indicates that Süskind "puts down the receiver"*, that they have been listening to a telephone conversation. Other cues are the one-sided nature of the conversation and that the scene takes place in Walter's office, since the AD mentions this as the location at the beginning of the scene. These cues are, however, much weaker in terms of cohesion than the visual input of the original clip, especially since the AD can only indicate briefly and in retrospect that the VIP has been listening to a telephone conversation. The linearity of verbal narration - or its specific textual potential, which is different from the potential of visual narration (that can show many objects and actions simultaneously) - limits the AD, and hence the VIPs' access to the multimodal redundancy of the original film.

However, not all occurrences of this type of dialogic interaction are necessarily problematic in the AD/AST version, as examples 21 and 22 demonstrate.

Example 21 (Süskind): In the continuation of the events discussed in example 19, Süskind is standing by the window. Having seen what is happening in the street below, he reacts with the words "Oh God" (in the dialogues and AST). His verbal reaction now follows the visuals of the German soldiers lining up the Jews in the street and this action is duly described by the AD. In other words, both the sighted and the VIP audience can rely on the functioning of the dialogic interaction between the film modes in order to interpret the scene.

Example 22 (Oorlogswinter): At the end of the scene transcribed in Table 2, the soldier who has opened the door of the City Hall and is speaking to Lia suddenly shouts, "Halt, blijf staan" [Stop, don't move] when Lia unexpectedly pushes him aside and bursts into the building. Again, it is the dialogic interaction between Lia's action and the soldier's line of dialogue that creates narrative cohesion. In this instance, too, the cohesion is fully recreated in the AD since "Halt blijf staan" [Stop, don't move] immediately follows the description: "Ineens duwt Lia de soldaat aan de kant en schiet naar binnen." [Suddenly Lia pushes the

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

70

soldier aside and forces her way into the building].

In none of the above examples does cohesion break down completely in the AD/AST version of the clips, and, as examples 21 and 22 show, the slot in which the AD is inserted can have an influence on the ease with which cohesion can be reconstructed in the accessible version of a film. In both examples, sufficient time was available for the AD to describe the action to which a character subsequently reacts verbally.

By contrast, in the other examples under discussion, the AD is limited by its textual potentiality (e.g. in example 20) and/or has to rely further on logical inference from the dialogue (e.g. the type of conversation in example 20) or other filmic modes (e.g. the sounds in example 19) because sufficient time is not available when it would be most appropriate to describe the action with which the dialogue interacts. In such instances our present analysis seems to suggest that the dialogic interaction of the different modes and the film's multimodal redundancy may not always be available for VIPs in the AD/AST version to the same extent as a sighted audience experiences. Consequently, a greater effort is required of VIPs to reconstruct the film narrative and re-create a fully coherent text. However, this is a preliminary analysis and a more in-depth multimodal analysis of the source text as well as reception studies with VIPs would be required to analyse these shifts and their overall effect on the textual cohesion of the accessible versions of the films.

6. Shifts in lexical and referential cohesion: food for thought

Our short analysis of the two clips from *Süskind* and *Oorlogswinter*, focusing on the way in which they create a coherent narrative combining AD and AST, shows that they make ample use of the cohesive cues identified by Halliday as referential and lexical allusion. One question that the analysis raises, however, concerns the extent to which the cues used in the AD and AST mimic the lexical and referential cohesion of the original clip. A further question worth investigating is the extent to which the somewhat different AST technique applied in each of the accessible versions has an influence on this cohesion.

Firstly, it must be underlined that in our particular samples both AST versions remain very close to the original German dialogue; hardly any rewriting or reduction of the spoken text

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

71

has been carried out. Secondly, and no doubt as a result of this, the analysis reveals that the cohesion between image and dialogue achieved through the use of reference in the original clip is almost entirely recreated in the versions combining AD and AST. Most of the referring items present in the AST, literal translations of the original dialogue, are picked up by the AD. Where the referring item in the original dialogue pointed to a visual element, the referring item in the AST now points to a described element in the AD, as demonstrated by examples 1 to 3.

There are, however, some instances, as shown in examples 6 to 9, where the referent of a referring item in the AST is implicit or even missing in the AD (while it was visually present in the original text). In those cases the overall coherence of the clip may be weaker in the AD/AST version as it is based more on inference than explicit visual textual cues. In example 7, for instance, the sighted audience can identify the places that the locative adverbs "hier" [here] and "daar" [there] refer to in the visuals, whereas the VIPs must infer where the Jews are hiding and create a mental image of the scene on the basis of the AST and the AD of the characters' actions.

Some of the other occurrences of an implicit referent are as implicit in the original clip as in the AD/AST version. In example 8, for instance, the sighted audience has no way of knowing what "het" [it] refers to in Walter's telephone conversation when he says, "Listen, Mr. Cohen, I assure you, it is a mystery to me too." Similarly, in examples 5 and 6, the references to characters who form part of the story but are not physically present in the scene (Süskind and the mayor, Van Beusekom, respectively), work only because the audience can activate their mental model of the narrative, in which these characters have already acquired a place. Whether the audience is using the original or the accessible version of the film may not make a difference in this case, since in both versions the cohesive link must be inferred by the audience, requiring a cognitive effort.

Thirdly, the analysis above shows that there is also some AD-AST cohesion through the reiteration of lexical items. Lexical cohesion is strongest in the case of repetition (example 7) but somewhat less so in the case of hyponymy or hyperonymy, since this means that the link between the two items is less explicit. However, it is difficult to evaluate to what extent such

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

72

lexical cohesion created in the AD/AST version can efficiently replace the cohesion created between visuals and dialogue. In the original clip, the referents of the lexical items in the dialogue are visual elements. In example 13 the referring item "mensen" [people], for instance, corresponds to the people we can see in the image, while in example 15 "de overkant" [the other side] corresponds to the façade of a building that is visible on screen. Due to the different nature of the visual and verbal modes it is virtually impossible to determine to what extent the dual multimodal reference "de overkant" [the other side] + the image of the façade is comparable in terms of cohesion to the dual verbal reference "de overkant" [AST/the other side] + "de crèche" [AD/the kindergarten]. Whatever the case may be, the AD attempts to recreate the cohesive link since the referents of the lexical items in the AST are picked up by the AD and rendered slightly more concrete in example 15. Whether this is sufficient to create a text that is as cohesively unambiguous and coherent as the original is something that can only be tested through reception studies.

Moreover, our analysis (see section 5) indicates that there is more to cohesion between AD and AST than the linguistic aspects of reference and lexical cohesion. Cohesion through different types of complementarity, which is implicit and relies on inference by the VIPs, plays an important role. Our preliminary analysis of a few examples also points to potential shifts in cohesion in the accessible versions due to their different handling of the narrative redundancy of the filmic medium (see section 5).

Another related issue that we have not yet investigated may also offer more food for thought. We did not investigate whether the type of recording used for the AST and the degree of adaptation of the spoken subtitles/translated dialogue have an influence on cohesion. When compared to the original dialogues, the AST translation as such does not seem to affect cohesion in our examples, since the translations in both our clips are quite literal. However, the different recording methods may have a certain impact on the overall sense of coherence. In the case of *Oorlogswinter*, the original German dialogue can still be heard before and underneath the AST. This means that certain words are repeated and that there is a different form of lexical (interlingual) repetition in the clip that might help recipients to create a more coherent story (see example 17). In contrast, in *Süskind* the German dialogues have been replaced entirely by the Dutch version.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

73

Example 17 (Oorlogswinter): When Lia says, "Goedemorgen, ik (möchte) zou graag (bitte) mijn echtgenoot willen spreken. Van Beusekom." [Good morning. I (möchte) would like to (bitte) speak to my husband. Van Beusekom]**,** for instance, the German words "möchte" [would like], and "bitte" [please] from the original dialogue, spoken by the same character and the same actor, are audible in between the AST (See also Table 2).

Dutch-speaking audiences are likely to know and understand simple German conversational words such as "bitte" and "möchte", which is also doubled up in "zou graag" [would like]. Moreover, since the users can still hear the original dialogue, they also know what the mother tongue of a particular character is, which helps them to identify speakers. This could be an advantage when trying to infer who might logically be saying what in a given context and to fit this information into the overall mental model the user has already created of the story. However, much depends on the clarity and timing of the voice-over recording, an issue that is beyond the scope of the present article (see Remael 2012 for a more in-depth analysis). In the case of *Süskind*, there is no interlingual lexical repetition where the original dialogue is substituted by a dubbed version. As a result, German soldiers speak perfect Dutch to each other and to others, which could certainly be confusing to VIPs.

7. Concluding remarks

This exploratory analysis of two film clips from the Dutch films *Süskind* and *Oorlogswinter* yields some interesting insights and invites more in-depth analysis of larger corpora of different types of films presenting different challenges.

Generally speaking, our analysis has revealed that films with AD and AST employ several "Hallidayan" cohesive devices that offer a potential for the audience to create a coherent story based on the accessible version of the film. The describers always try to ensure that the referring items in the AST have some corresponding referent in the AD. As a result, the internal textual cohesion remains intact in the accessible version of the film. In fact, reference and lexical cohesion work closely together to produce textual cohesion as they create cohesive threads across the AD and AST text (see examples 2 and 11). Having said that, it seems implicitness and inference also play a significant role in ensuring cohesion and coherence in multimodal texts. Some examples indicate that considerable effort is required by

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

74

an audience to reconstruct a coherent text on the basis of the cohesive cues in the AD/AST version. This prompts the question as to whether the cohesion in the accessible version is in fact as strong as the cohesion of the original clip or whether a greater cognitive effort is required of the audience. Some of the examples discussed above point towards the presence of what we have called cohesive shifts in the translation process.

Some of these apparent shifts involve the substitution of verbal-visual cohesion with a form of verbal-verbal cohesion involving hyponyms, as (e.g. "iemand" [someone] and "een vrouw" [a woman]) which might not be as explicit as the original visual input. In addition, it might be quite challenging to link pronouns such as "hier" [here] or "daar" [there] to a concrete location if the AD describing or suggesting the location follows rather than precedes the dialogue. In other cases, the referent to which a referring item refers may be implicit or omitted in a given scene, but will have been mentioned at an earlier point in the narrative, thus requiring the VIPs to draw on their mental model of the story world to retrieve the cohesive cue. Even if the referent is omitted for sighted and VIP audiences alike and cannot be derived from information in an earlier scene, it may be easier to infer what or who the referent of a conversation is, or how important or unimportant it is when the context in which the conversation takes place is visually present on the screen (e.g. "het" [it] in Süskind's telephone conversation).

Likewise, and particularly relevant for AVT and Media Accessibility in general, is the cohesion created through complementarity, which seems to play an important role in multimodal cohesion. However, at times this type of cohesion seems to be weaker in the AD/AST version, because of the limitations imposed on the AD by the different potentiality of the verbal text, and more specifically, its linear nature. This can, indeed, result in less scope for multimodal redundancy because images can show more in one shot than words can tell. Loss of redundancy, especially the need to be more reliant on film sound when the time available is insufficient to describe the action, may also impact on the degree of cohesion. This confirms the need for research into recording techniques that do justice to all the elements of the sound track on the one hand, and the need for AD scripts to avoid covering sound effects on the other. Our examples also seem to confirm that a greater degree of explicitness may be required when AD/AST versions resort to different forms of cross-

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

75

referencing (see Orero and Braun 2010), and that it may be better to give preference to the AD preceding a verbal reaction rather than following it. These preliminary hypotheses regarding the greater effort that may be involved in processing the AD/AST version of a film are an interesting avenue for further research identified by our analysis. This requires more in-depth analysis through detailed multimodal analysis and transcription of both the ST and the TT, and reception research with blind and visually impaired patrons.

On a more methodological level, the Hallidayan concepts detailed in the introduction (section 1), more specifically, reference and lexical cohesion, prove to be useful tools for the discussion of the linguistic cohesive cues used to create coherence in accessible films that combine AD with AST. Nevertheless, the differences evident in no more than two film clips beg further research with more clips from different genres. Moreover, some forms of multimodal cohesion have been shown to rely on a type of complementarity and the background to how it functions cannot be captured by linguistic analysis alone since in these cases the interaction between visual and non-visual filmic modes is crucial. Our examples confirm that it is essential for AD/AST to account for the interaction between dialogue and visual action/events in films, as Braun (2007) and Yos (2005) point out. However, the different success rate in the degree of coherence achieved in our examples again suggests that the analysis of a greater number of films or clips from different genres is required. Other films will no doubt uncover additional challenges for films combining AD and AST with respect to the recreation of multimodal redundancy and linguistic cohesion.

References

Baldry, Anthony (2005) *A multimodal approach to text studies in English. The role of MCA in multimodal concordancing and multimodal corpus linguistics*, Campobasso: Palladino Editore.

Baumgarten, Nicole (2008) 'Yeah: That's it!: Verbal reference to visual information in film texts and film translations', *Meta: Translators' Journal,* 53(1): 6-25.

Bordwell, David and Kristin Thompson (2010) *Film art: An introduction* (9th ed.), New York: McGraw-Hill.

Braun, Sabine (2011) 'Creating coherence in audio description', *Meta: Translators' Journal,* 56(3): 645-662.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

76

Braun, Sabine and Pilar Orero (2010) 'Audio description with audio subtitling – An emergent modality of audiovisual localisation', *Perspectives: Studies in Translatology,* 18(3): 173-188.

Chaume, Frederic (2002) 'Models of research in audiovisual translation', *Babel,* 48(1): 1-13.

Delabastita, Dirk (1989) 'Translation and mass-communication', *Babel*, 35(4): 193-218.

Díaz Cintas, Jorge (2011) 'Dealing with multilingual films in audiovisual translation', in I. Ohnheiser and W. Pöckl (eds) *Translation, Sprachvariation, Mehrsprachichkeit*, Frankfurt: Peter Lang, 215-233.

Emmott, Catherine and Marc Alexander (2009) 'Schemata', in P. Hühn, J. Pier, W. Schmid, J. Schönert (eds) *Handbook of Narratology,* Berlin: de Gruyter, 411-419.

Gambier, Yves (2013) 'The position of audiovisual translation studies', in C. M. Carmen Millan-Varela and F. Bartrina (eds), *The Routledge Handbook of Translation Studies*, New York: Routledge, 45-59.

Halliday, M.A.K and Rukiyan Hasan (1976) *Cohesion in English,* London: Longman.

Johnson-Laird, Philip (1983) *Mental models: Towards a cognitive science of language, inference, and consciousness,* Cambridge, MA: Harvard University Press.

Kress, Gunther and Theo Van Leeuwen (1996) *Reading images: The grammar of visual design*, London/New York: Routledge.

Marleau, Lucien (1982) 'Les sous-titres…un mal nécessaire', *Meta: Translators' Journal,* 27(3): 271-285.

Norris, Sigrid (2004) *Analyzing multimodal interaction: A methodological framework*, London/New York: Routledge.

O'Sullivan, Carol (2011) *Translating popular film*, Basingstoke: Palgrave Macmillan.

Remael, Aline (2012) 'Audio description with audio subtitling for Dutch multilingual films: Manipulating textual cohesion on different levels', *Meta: Translators' Journal*, 57(2): 385-407.

Remael, Aline (2001) 'Some thoughts on the study of multimodal and multimedia translation', in Yves Gambier and Henrik Gottlieb (eds.), *(Multi)media translation,* Amsterdam/Philadelphia: John Benjamins,13-22.

Royce, Terry D. (2007) 'Intersemiotic complementarity: A framework for multimodal discourse analysis', in Terry D. Royce and W.L. Bowcher (eds.), *New Directions in the Analysis of Multimodal Discourse*, Mahwah, NJ: Lawrence Earlbaum, 63-109.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description: A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

77

Serban, Adriana and Reine Meylaerts (2014) (eds), 'Multilingualism at the cinema and on stage: A translation perspective', *Linguistica Antverpiensia NS 13 - Themes in translation studies* (https://lans-tts.uantwerpen.be/index.php/LANS-TTS/issue/current).

Sindoni, Maria Grazia (2011) *Systemic-Functional Grammar and Multimodal Studies: An Introduction with Text Analysis*, Pavia: Ibis.

Tincheva, Nelly (2012) *Enter Text Linguistics*. Sofia: Direct Services Publ.

Vercauteren, Gert and Aline Remael (2014) 'The audio-description of spatio-temporal settings', in Anna Maszerowska, Anna Matamala and Pilar Orero (eds), *Audio Description: New Perspectives Illustrated,* London: John Benjamins, 61-80.

Yos, G. (2005) 'Verknüpfungen von Audiodeskription und Filmdialog', in Ulla Fix (ed.), *Hörfilm: Bildkompensation durch Sprache,* Berlin: Erich Schmidt Verlag, 99-116.

Zabalbeascoa, Patrick (2008) 'The nature of the audiovisual text and its parameters', in Jorge Díaz Cintas (ed.), *The Didactics of Audiovisual Translation*, London: John Benjamins, 21-38.

Filmography

Desmyter, Stef (2012-2013). *Wolven* [Wolves]. Belgium

Koolhoven, Martin (2008). *Oorlogswinter* [Winter in Wartime]. The Netherlands.

Runze, Ottokar (1994). *Tatort. Laura, Mein Engel* [Crime Scene. Laura, my Angel]. Germany.

Van den Berg, Rudolf (2010). *Tirza* [Tirza]. The Netherlands.

Van den Berg, Rudolf (2012). *Süskind* [Süskind]. The Netherlands.

Van Looy, Eric (2003). *De Zaak Alzheimer* [The Alzeheimer Case]. Belgium.

Verheyen, Jan (2009). *Dossier K* [The K Files]. Belgium.

Webber, Peter (2003). *The Girl with the Pearl Earring*. USA.

Nina Reviers and Aline Remael, Recreating Multimodal Cohesion in Audio Description:
A Case Study of Audio Subtitling in Dutch Multilingual Films, 50-78.

78