

Contrasting Visual and Verbal Cueing of Space: strategies and devices in the audio description of film

Maija Hirvonen

University of Helsinki, Finland

ABSTRACT

This article studies audio description, a form of intermodal (or intersemiotic) translation that renders visual and, occasionally, auditory information in a verbal-vocal form for the benefit of visually impaired people. It compares visual and verbal representation in the context of narrative film in which space is strongly cued through the visual channel. The focus is on shot distance, which controls the amount of space visible in a shot. The study analyses how shot distance is reflected in audio description by syntactic and semantic means. Four different-language audio descriptions of two films were analysed, contrasting the visual source text with the verbal translation. The study aims to show how audio description can make use of diverse representational strategies and linguistic devices in rendering shot distance. These strategies and devices could be used purposely to compensate for visual cues so as to give an idea of space similar to that conveyed by the visual representation.

KEYWORDS: audio description, film, intermodal translation, linguistic devices, representational strategies, shot distance, spatial cues.

Introduction

Audio description (AD) is a form of translation that renders visual and, occasionally, nonverbal auditory information in a verbal-vocal form for the benefit of visually impaired people. By providing verbalizations of the visible and by explaining sources of sounds, AD supports the comprehension of audiovisual or visual communication, art and entertainment such as film, theatre and museum exhibitions. For blind people and those with severe loss of sight, AD is a capacitating aid that renders the visual world accessible; and for people with milder degrees of low vision, it can function as an additional support which enhances reception. In film, AD renders visual narrative cues (and sometimes sources of sounds) through language, delivering descriptions in a spoken form in the silent gaps between dialogues and important sound effects. AD contributes to the social and cultural integration of visually impaired persons by facilitating an integral and independent reception of films.

In translation studies, AD is receiving growing attention. Some of the earliest topics include descriptions of AD as a translation phenomenon (see e.g. Benecke 2004; Hyks 2005; Snyder 2005) and descriptions of the social context in which AD is carried out (e.g. Orero 2007). Two edited books (Fix 2005; Diaz-Cintas et al. 2007) and a more recent special issue (Kruger and Orero 2010) present a variety of research topics, including didactic, linguistic, narratological, and translational aspects. AD has inspired analyses contrasting, on the one hand, source and target texts (e.g. Benecke 2007; Remael and Vercauteren 2007) and, on the other hand, parallel target texts (Bourne and Jiménez 2007; Matamala and Rami 2009). However, research on the meaning transfer that is at the core of AD is still lacking (Braun 2008:16), particularly from the point of view of the source text (e.g. a film). The present article sets out to address this gap, studying filmic AD from the source-text perspective: that is, contrasting the source text's visual representation with the corresponding audio description.

More specifically, the article focuses on one technique used in visual film narration, namely shot distance, and its verbalization in AD. While film can imitate the visual world directly through visual cues, audio describers must select and verbalize what aspects they describe and how they do it. Shot distance is a technique of spatial cueing that controls the amount of space visible in a shot by varying the distance of the camera from the object that is filmed. It is particularly effective when a new scene begins as a means to control the viewers' first impression of a particular space (e.g. offering a panoramic view of the setting or a close-up of a character). This study examines how AD may compensate for the functions of shot distance. It analyses the types of shot distance involved in the scene openings of two narrative films (*Der Untergang* [Downfall], 2004, and *Dancer in the Dark*, 2000) and in their four different corresponding audio descriptions (German and Spanish for the first film, English and German for the latter). The aim of the study is to define representational strategies and linguistic devices that can be used in verbalizing the visual cues provided by shot distance.

The research has both practical and theoretical implications. It explains how one visual cue is translated into a linguistic representation and it is hoped that it will inspire practical work by presenting strategies and devices for AD. Furthermore, the findings may provide new insights into the concept of similarity in translation (Chesterman 1996; 1997) by studying the possibilities of similarity in intermodal translation.¹ If one chooses a "fidelity" approach to audio-descriptive translation, then shot distance could be considered as a relevant feature of the filmic source text that can be rendered in AD (cf. Nord 1991:22).

The article is organized as follows: the first section deals with the theoretical background, shedding light on the process of filmic AD and the use of shot distance as spatial cue in film. The second section describes the research questions and objectives, data, and methodology in greater detail. The third section demonstrates the qualitative analysis that was carried out in order to define AD strategies and devices and the findings of this analysis are then reviewed with a quantitative approach in the fourth section. The last section outlines the most important conclusions and offers a brief discussion of possible implications of the study.

1. Theoretical background: filmic audio description and shot distance

1.1 Filmic audio description: process and factors

Audio description has been variously termed intersemiotic, intermodal or cross-modal translation or mediation (Braun 2008:15-16). The term adopted in this article is 'intermodal' and is based on a comprehensive concept of 'mode' as both a semiotic resource and a modality of communication (cf. Kress 2009/2011). Thus, 'intermodal' characterizes audio description in the sense that it is translation between semiotic resources (visual images and language) and between modalities (visual and auditory). Audio describing an audiovisual text such as film involves a multilayered process of (a) *intermodal translation* from the visual (and, occasionally, auditory) meaning-making system into verbal language, so that verbal descriptions of the visual modality are rendered audible through speech, and (b) *multimodal mediation* of the verbal descriptions in suitable slots, i.e. dialogue-free passages, in the original soundtrack to form synergetic communication of sound, music, and language (cf. Braun 2008:15-16).

¹ I have chosen to use the concept of 'similarity' instead of 'equivalence' for two reasons: firstly, equivalence is a controversial issue in translation studies overall and, secondly, it can be considered as too narrow in scope to deal with an area of translation that is not interlingual.

For pre-recorded AD (e.g. for a cinema screening or DVD), the process typically runs as follows (Orero 2007:116-117): selected visual and auditory elements are verbalized and adjusted according to the original soundtrack into a written script. Based on the script, the verbalizations are narrated by an actor (or by the audio describer himself/herself) and recorded as a separate sound. Finally, the audio description is mixed to the original filmic soundtrack either in a studio or on the spot in the cinema.

The process involves a variety of factors that affect both the translation and mediation, including, time and timing; technical implementation; multimodality; and intermodality.

a) *Time and timing*: Since AD should not interfere with important auditory narration, the descriptions are generally bound to the gaps between dialogue and sound effects. The time available for description varies according to the original soundtrack and may, in some cases, consist of only one or two seconds. Timing is another factor (Braun 2008:17); AD cannot always be provided simultaneously with the information to which it refers (e.g. in the case of coinciding dialogue), so that descriptions may be delayed or given in advance.

b) *Technical implementation*: Filmic AD, like any form of audiovisual translation, involves several technical procedures, from using editing software in the translation process to the need for sound engineering in the mediation phase. In delivery and transmission, quality of sound is central from the standpoint of reception (for example, poor sound quality may impede understanding).

c) *Multimodality*: The filmic soundtrack plays a role in the overall reception of an audio described film in the sense that sounds, music and dialogue communicate meaning in conjunction with the verbal descriptions (e.g. the sound of an opening door may be followed by a description of the person who enters the room) (cf. Fryer 2010). These multimodal ways of meaning making must be considered in the process of writing the AD script and mediating it vocally (Braun 2008:16).

d) *Intermodality*: The factor of intermodality arises from the translation process itself, that is, from the transfer between modes. The central question here is just how much can or cannot be done with AD. Some of the differences between visual and verbal communication may be outlined according to Stöckl (2004): visual communication represents objects holistically and simultaneously, offering abundant information at a glance, and also iconically, so that, for instance, photographic images show the represented object directly. The verbal mode is linear, offering pieces of information in segments, and refers to objects conceptually instead of showing them (ibid.). Nonetheless, visual and verbal modes seem to converge in some respects. One aspect is iconicity in language, or the extralinguistic motivation for linguistic expressions, which can be produced through sequential ordering (see e.g. Haspelmath 2008): word order reflects the order in which things are perceived. This is relevant for AD in the sense that the linear structure of the description can be made to reflect the order in which objects are represented in the visual mode.

1.2 Shot distance as a spatial cue

Film narration makes use of visual and auditory cues, on the basis of which spectators create mental representations and interpret the narrative (Bordwell 1985:99ff.). One such cue is shot distance. Shot distance (Monaco 2009:221), or camera distance in the terminology of Bordwell and Thompson (1990:176), refers to the “sense of distance” the spectator has in relation to what is visible in the shot (ibid.) or to “the amount of subject viewed” (Monaco

2009:221). A shot that frames a character's face in a close-up certainly gives an impression distinct from one showing a character's whole body or panoramic scenery. Shot distance is a strong spatial cue, particularly in scene openings where the opening shot, i.e. the first shot of the scene, quickly cues the first spatial impression. Films make use of different types of shot distance, each of which have specific functions and characteristics. These are described in Table 1, a synthesis from Bordwell and Thompson (1990), Straßner (2001) and Monaco (2009).

Table 1: Descriptions of standard shot distances

Shot distance	Abbr.	Description
Extreme long shot	ELS	Setting is dominant; human figure barely visible (attracts attention when mobile)
Long shot	LS	Human figure(s) becomes more prominent but setting still dominates
Medium long shot	MLS	Human figure(s) is framed from knees up, (inter)action becomes dominant
Medium close-up	MCU	Human body from chest up foregrounding interaction, gestures and facial expressions
Close-up	CU	Part of body or a small object is framed, emphasizing facial expression or detail
Extreme close-up	ECU	Portion of a face or object is visible, isolating a detail

Levin and Wang (2009:36) use the categories of long, medium and close coverage to describe how much space is visible in a shot. Following this idea, it is possible to define three strategies for spatial cueing, respectively driven by setting, interaction and detail.

The *setting-driven strategy* is realized when an extreme long shot (ELS) or a long shot (LS) depicts a long view of the scene. Thus, the primary focus of attention is on the setting. The ELS is typically used at the beginning of a scene (i.e. a narrative continuum) as an establishing shot, showing the general environment or situation (Bordwell and Thompson 1990:121). The LS also depicts the setting, but characters and (inter)action are more prominent than in the ELS. These shot distances “connect a character and a location in a single shot” (Katz 1991:128).

The *interaction-driven strategy* means using a medium long shot (MLS) or a medium close-up (MCU) to focus on the action of characters and interaction between them. In interaction-driven shots, the amount of space viewed ranges from showing one or more characters (in the MLS) to depicting a single character (in the MCU). The closer the shot is to the characters, the more their gestures and facial expressions are emphasized (Bordwell and Thompson 1990:176). Since they are focused on characters, interaction-driven shots often include dialogue (cf. Straßner 2001:1095).

The *detail-driven strategy* shows a detail in space through a close-up (CU) or an extreme close-up (ECU). The CU shows a face or body part, and the ECU covers merely a portion of a

face or a small object. Detail-driven shots are typically used to focus on facial expressions or specific objects in the setting.

This typology (setting-, interaction-, and detail-driven strategies) offers a generalized description of shot distance as a spatial cue that, as will be shown in the analysis, may be applied to the study of AD. Although the boundaries of shot distance functions are fuzzy (e.g. interaction may be the salient element in setting-driven shots) and other narrative devices affect the experience of space (e.g. sound or light may draw attention to a small object within a panoramic shot), the typology points to some basic differences in shot distances and determines one relevant feature of film narration, namely, the cueing of space in the sense of *coverage*, the amount of space viewed (cf. Levin and Wang 2009:36). The present research holds this spatial cue to be a feature of the source text that can be found in the target text with the aim of describing and understanding translation solutions (cf. Nord 1991).

2. The study: contrasting shot distance with audio description

2.1 Research questions and objectives

Focusing on shot distance as a spatial cue, this study addresses the following questions. Firstly, how much space is cued in the source text (film), i.e. what is the amount of space visible through shot distance in an opening shot, and secondly, how much space is cued in the translation (AD), i.e. what is the amount of space that is referred to in the audio descriptions, and what is its relation of similarity to the source text?

Contrasting the visual representation in the source text to the linguistic cueing in the target text, the study aims to describe representational strategies and linguistic devices that can be deployed in AD to verbally cue spatial aspects in the sense of setting-, interaction- and detail-driven shots. Consequently, the analysis seeks to define similarity between the *spatial coverage* in film and the *referential coverage* in AD.

2.2 Preceding studies

The present study is inspired by the analyses in Kluckhohn (2005), Seiffert (2005) and Braun (2007). They have provided interesting findings with regard to the strategies of representational similarity and linguistic devices in AD.

Kluckhohn (2005) analyses information structure as a compensation strategy in AD. He shows that linguistic information can be “packed” in a way that is optimally processed by the listener and that these choices may be used to compensate for the visual discourse in film. For example, manipulations in word order can help to clarify scene changes (Kluckhohn 2005:58-62). Seiffert (2005) analyses how words trigger knowledge and emotions in the mind. For instance in German, *Raum* (‘room’ or ‘space’) is more abstract than *Büro* (‘office’), and on hearing the word *Büro*, a specific schema (knowledge representation in the mind) is activated in the listener, and knowledge becomes available (Seiffert 2005:77). Seiffert concludes that it seems possible in AD to re-construct filmic space in the mind of the listener by using keywords that activate appropriate schemata (2005:83-84).

Braun discusses two strategies in creating a verbal account of a film’s visual cues (2007:8): AD can refer explicitly to the mental representations and associations the cues trigger (e.g. a shot of a woman writing at a desk described as “she sits writing a letter/a diary”), or it can verbalize no more than the individual cues in order to leave scope for individual interpretation (e.g. the same shot described as “she sits writing at a desk”). Whereas the studies by

Kluckhohn (2005) and Seiffert (2005) begin from the viewpoint of target text (a German audio-described film), the present research starts from the source text and compares it to the translation. A source-text-based approach has been adopted by Braun (2007) as well as by Pérez Payá (2007), but their analyses remain exploratory.

2.3 Data and methodology

The research material consists of opening shots and opening audio descriptions gathered from two films and four different audio descriptions (Table 2).

Table 2: Research corpus

	Film Title	Year	Director	Language	AD
1.	<i>Der Untergang</i>	2004	Oliver Hirschbiegel	German	German
	<i>El hundimiento</i>			Spanish (dubbed)	Spanish
2.	<i>Dancer in the Dark</i>	2000	Lars von Trier	English	English
	<i>Dancer in the Dark</i>			German (dubbed)	German

Der Untergang (henceforth *Untergang*) [*Downfall*] is a war drama that narrates the last days of Hitler in Berlin at the end of World War II; the story is based on the memoirs of his private secretary Traudl Junge, who also appears in the film as a main character. The film has been audio described in German (in the original version) and in Spanish (in the dubbed version). *Dancer in the Dark* (henceforth *Dancer*) is a musical drama about an immigrant single mother in the USA of the 1960s, starring the famous Icelandic singer Björk. *Dancer* has been audio described in English (original) and in German (dubbed). These films were chosen as the research material because, in 2008 when I was looking for films that entail more than one audio description in languages that I study (English, Finnish, German and Spanish) and that were available to purchase on DVD or of which I managed to get a DVD copy, I was only able to find these two films.² Thus, since *Untergang* and *Dancer* have been audio described in more than one version, they make a good case for comparison: one source text is contrasted to two different translations. However, it should be noted that this study is not cross-linguistic and does not contrast language-specific ways of audio describing but rather looks at language as a general resource for meaning making as opposed to visual imagery and, therefore, intends to define general strategies and devices for linguistic representation.

The opening shots and the opening audio descriptions that are the focus of this study all occur in scene transitions in which the narrative changes so that the central action and/or the setting are transformed. As Kluckhohn has suggested, transitions may be particularly challenging for AD because a potentially large amount of new information must be given in order to introduce the new situation (2005:63-64). Notwithstanding this, viewers can perceive this

² Since 2008, more films have been released on DVD with different-language audio descriptions: among others, *Happy Go Lucky* (Mike Leight, 2008; audio described in English in 2008 and in German in 2009) and *Slumdog Millionaire* (Danny Boyle & Loveleen Tandan, 2008; audio described in German, English and Spanish in 2009). Currently, several online sources provide information about audio-described films in individual countries, such as 'Your Local Cinema' in the UK (<http://www.yourlocalcinema.com/ad.dvd.html>, accessed 27 March 2012), 'Hörfilm e.V.' in Germany (<http://hoerfilmev.de/index.php?id=117>, accessed 27 March 2012), and 'Fundación Orange' in Spain (http://fundacionorange.es/fundacionorange/proyectos/proyecto_cineaccesible_dvd.html, accessed 27 March 2012).

change rapidly (through the visual channel), recognizing (and identifying) the new space through visual cues, such as objects and characters. Sometimes the new situation is explicitly visible as in setting-driven shots. At other times, information about the new scene may be delayed or obscured (cf. Seiffert 2005:84-85), for example by using a detail-driven opening shot.

Following the method of Nord's translation-oriented text analysis (1991), in which relevant features of the source text are sought in the target text, and Chesterman's take on translation strategies as also observable in the translation product (manifested as textual manipulation) (1997:89), the present study used three steps to analyse the shot distances of opening shots and their corresponding audio descriptions.

Firstly, the audiovisual material was described in textual form in order to perform verbal searches (cf. Morgner and Pappert 2005). Then, the first stage involved locating scene transitions in the data using also the film scripts as source material (Trier 2000; Fest and Eichinger 2005). The number of scene transitions and, hence, opening shots, came to 166 in *Untergang* and 75 in *Dancer*. Shot distance in the opening shots was analysed using the typology presented in the theoretical background section 1.2.

The second stage involved locating the source text feature in the translations through an analysis of the verbalization in the opening descriptions, termed here *AD openings*. The first utterance of each new scene was analysed in terms of linearization (how the new space is cued structurally) and referencing (what kind of words and expressions are used). The "beginning point" of each new scene, the first word or expression, was considered as a central cue because (i) it constitutes the initial thematic context for the text that follows (cf. Brown and Yule 1983:125) and (ii) the sentence-initial position is regarded as cognitively prominent, i.e. it receives the most attention in an utterance (Lambrecht 1994:51, 201). Reference is "treated as an action on the part of the speaker/writer" (Brown and Yule 1983:28) to make an intended referent (a visually perceivable object) mentally accessible and identifiable for the receiver (Blakemore 1992:68-69). For the intermodal comparison, linguistic cueing (linearization and referencing) was examined in the light of setting-, interaction- and detail-driven scene openings of the films, analysing the way in which the linguistic cues draw attention either to the setting, to interaction or to a detail in the scene.

In the third stage, this comparative analysis was followed by a quantitative analysis that resulted in a systematic account of strategies and devices. To this end, different AD openings were classified according to the representational strategy and the linguistic device they manifest, and their occurrence in the data was counted. Both analyses will be presented next in sections 3 and 4 respectively.

3. The comparative analysis

This section illustrates the comparative analysis through selected examples that manifest instances of similarity in the three categories of spatial coverage: setting-driven, interaction-driven and detail-driven openings. Since it was not possible to use the original still images,³ shots are reconstructed as schematic sketches depicting the shots' spatial organization and thereby suggesting the shot distance in the film. Corresponding AD openings are presented as transcriptions and with the scene transition in which they occur, including: the last description

³ Using original still images of *Untergang* is subject to a charge that I was not prepared to pay. Thus, I decided to treat the visual data of both films equally and reproduce the shots with sketches.

and/or sound effect from the previous scene, the AD opening (highlighted in bold font), and the subsequent utterance and/or sound effect. Since I want to illustrate how AD is received by the audience, I have selected GAT(2) ('Gesprächsanalytisches Transkriptionssystem 2' [system for transcribing talk-in-interaction 2]) from Selting et al. (2009) as the transcription format to represent AD as spoken interaction within its filmic, auditory context; this involves marking prosodic features such as pauses and other auditory but nonverbal actions such as sound effects. GAT(2) is a conventional form of representing spoken language and interaction and is also deployed to represent AD by Morgner & Pappert (2005) and Hirvonen & Tiittula (in press). One characteristic feature that might strike readers familiar with the German orthography should be mentioned: the GAT(2) system requires the non-capitalization of names as well as German nouns because it reserves capital letters for illustrating word or syllable stress. Following the above-mentioned studies, I will also use Courier as the transcription font. A legend of the transcription symbols is provided in the appendix. The different audio descriptions are codified as follows: AD-DE for German audio description, AD-ES for Spanish, and AD-EN for English.

3.1 Setting-driven scene openings

3.1.1 Focusing on the setting

Image 1 illustrates a scene transition in *Untergang* in which the new scene opens with a setting-driven shot, in this case, an extreme long shot.

Image 1: Scene 19/*Untergang*



The last shot of scene 18 is a medium close-up of Burgdorf, Hitler's general, who is having a discussion with other officers inside Hitler's bunker. The next shot (numbered as 19a) is an establishing shot that opens scene 19. The transition between scenes 18 and 19 changes the view from the character to a large amount of shot space, thus focusing on the setting: the shot shows a beaten-up courtyard, a white building in the background plane and a dark gray cement cubicle in the middle plane. The auditory cue is distant explosions and machinegun fire, which seem to originate from off-screen space (a war is being fought in the city). After a few seconds, movement can be perceived at the cubicle as characters, recognizable as soldiers, come out of the cubicle door and begin to gather around the tree trunk beside it (illustrated as the small figures in the middle of 19a). This opening shot lasts approximately 6 seconds, thus leaving time for the viewer to explore it. The next shot (19b) is a long shot that allows a closer view of the cubicle and the action that takes place beside it.

Extracts 1 and 2 present the two opening descriptions of this scene in the German and the Spanish version respectively, and illustrate how the setting as relevant spatial cue is emphasized by the audio description. Back-translations into English are provided in italics below the transcription lines. The scenes are marked in parentheses on the left.

Extract 1: Scene 19/AD-DE/*Untergang*
(Scene 18)

AD: fegelein- starrt ins leere.
Fegelein stares into space

(Scene 19) [distant explosions + machinegun fire]
im verwüsteten garten hinter der alten reichskanzlei.
in the devastated garden behind the old Reich Chancellery
(.) rauch hängt in der luft. (-)
(.) *smoke is hanging in the air*

Extract 2: Scene 19/AD-ES/*Untergang*
(Scene 18)

FE: qué ha dicho?
what did you say?

AD: **fuera del búnker los soldados**
outside the bunker the soldiers

(Scene 19) [distant explosions + machinegun fire]
se cuadran ante la salida de hitler;
are standing to attention as Hitler exits (the bunker)
un oficial saluda brazo en alto.
an officer salutes with a raised arm
((bombing))

In the AD-DE (Extract 1), the new scene is first cued by ‘in the devastated garden behind the Old Reich Chancellery’,⁴ with the sound of gunfire in the background. The AD-ES (Extract 2), on the other hand, begins to describe the new scene already at the end of the previous shot: ‘outside the bunker the soldiers...’. This utterance continues as the soundscape of the new scene plays in the background: ‘...are standing to attention as Hitler exits (the bunker)’.

Both openings mark a new topic and, thus, a narrative transition by structural emphasis: focus construction and topicalization. Focus construction - the term is adapted from Kluckhohn’s *All-Fokus-Konstruktion* (2005:53) - is an elliptical utterance that indicates a change of scene and introduces a new situation: “im verwüsteten garten hinter der alten reichskanzlei” (‘in the devastated...’). Topicalization, too, can be used to mark a transition (Kluckhohn 2005:61), and it modifies a sentence syntactically so that a non-subject item is foregrounded by placing it in a cognitively prominent position (the sentence-initial position) (Lambrecht 1994): “fuera del búnker...” (‘outside the bunker...’). When used as the beginning point for a new situation, the items rendered through focus construction and topicalization put the referent in our attentional foreground. In both openings, the new topic is a setting: in the AD-DE, a devastated garden behind the old Reich Chancellery, and in the AD-ES, an area around the bunker. The Spanish audio description simulates the possibility of setting-driven shots to shift viewers’ attention in space by referring first to setting and then to characters and their action; the AD-DE, in contrast, remains with the setting. Moreover, the audio descriptions direct attention to distinct entities. The AD-ES frames the shot space from the cement cubicle by referring to the bunker, whereas the AD-DE verbalizes its surroundings, i.e. the garden and the building in the background. A long view thus seems to provide for variation in selecting the frame of reference for AD.

⁴ In this analysis, I will refer to the data in the following ways: the original utterances, or quotes from the transcriptions, are rendered in double quotations (“”) and back-translations in single quotations.

3.1.2 Focusing on a setting element

The next example (Image 2) is from *Dancer* and illustrates the kind of opening shot in which a particular condition of light dominates the image and may impede or delay the recognition of the setting.

Image 2: Scene 31/*Dancer*



Scene 30 shows a dialogue between the protagonist Selma, her friend Kathy and her suitor Jeff in the factory car park as Selma and Kathy are leaving at the end of their night shift. Jeff, who is attracted to Selma, has come to offer Selma a ride home but she refuses and, after a short dialogue, starts to walk towards the railway tracks that lead to her home. Jeff and Kathy, who is worried about Selma because she is losing her sight, watch as she slowly makes her way. The scene ends in a close-up of the worried-looking Kathy (30). The soundscape is silent. (31a) reconstructs the opening shot of scene 31, illustrating how a physical element may dominate the shot space: here it is the lack of light depicted by the black and dark gray areas in the image. The white in the image illustrates a brighter area that is the sky. The white lines and the dark figure in the middle represent objects that stand out in the setting: railway tracks and the figure of a walking character, who may soon be recognized as Selma, in the middle of the frame. In the subsequent shot (31b), the view of the character is closer.

The lack of light is reflected in the audio descriptions as both the AD-DE and the AD-EN lay emphasis on the illumination (Extracts 3 and 4).

Extract 3: Scene 31/AD-DE/*Dancer*

(Scene 30)

AD: kathy sieht ihr verzweifelt nach.
Kathy follows her with her eyes in despair

KA: ((gasps))

(Scene 31)

AD: **es dämmert.** (--)
it is getting light (--)
über der bahnstrecke liegt schimmernder nebel. (-)
over the railroad track lies shimmery fog

Extract 4: Scene 31/AD-EN/*Dancer*

(Scene 30)

AD: with tears welling in her eyes (.) kathy turns away
(.) her hand to her mouth. (--)

(Scene 31) **in the gray light of dawn-** (.)
selma walks steadily along the track. (--)
[bird sings]
little wisps of mist cling to the trees.

In the German audio description (Extract 3), “es dümmert” cues the first impression of the new scene as a yet unidentified space with a certain type of light that indicates twilight or dawn; the German verb *dümmern* does not specify which time of day is meant. In the current context, it may be interpreted as dawn (‘it is getting light’) because the preceding scene happened at night. The German pronoun *es* (‘it’) is used as zero subject to profile a concrete, physical setting or meteorological condition (see also the analyses in Smith 2002 and, with reference to audio description, in Hirvonen and Tiittula (in press)). In the English parallel (Extract 4), the topicalized expression “in the gray light of dawn” opens the same scene also by referring to the illumination. This description, however, renders the visual cue of “gray light” more explicitly, i.e. as “dawn”, which remains implicit in the German description “es dümmert”.

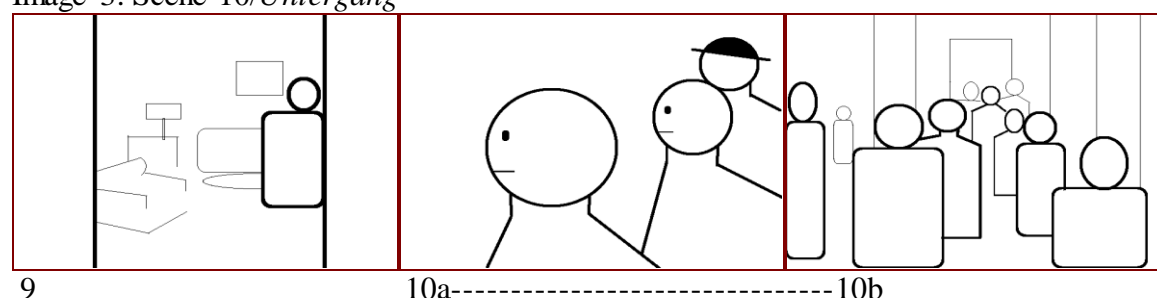
Although a similar construction that profiles a physical condition in the setting is possible in English (e.g. *it is raining*) as well as in Spanish (with zero pronoun, e.g. *Ø está lloviendo* (‘Ø is raining’)), the English and Spanish audio descriptions in the data deployed other means for opening scenes characterized by illumination: typically they use focus construction, for example, “night” (in an interaction-driven opening as in 11/AD-EN/*Dancer*⁵) or topicalization, as in “de día la gente camina entre los escombros” (‘in the daytime people are walking among the debris’) (in a setting-driven opening as in 75/AD-ES/*Untergang*⁶).

Not only in setting-driven shots but also within other types of shot distance, illumination can be a salient element, as the example from Scene 31 in *Dancer* shows. Furthermore, it is not merely a visual, aesthetic element of shot space but also cues temporal transitions of the narrative. AD can relate to both functions in different ways, as our examples demonstrate, thus describing the illumination in the scene (Extracts 3 and 4) or referring more directly to time (‘night’; ‘de día’).

3.2 Interaction-driven scene openings

The second part of the qualitative analysis deals with interaction-driven openings, which focus on characters and their action. Image 3 shows such a transition from *Untergang*.

Image 3: Scene 10/*Untergang*



⁵ In Scene 11 of *Dancer*, interaction that happened in an interior setting in the previous scene changes to an exterior setting. The first shot shows Selma and Linda, Selma’s landlady and neighbour, exchanging a few words at Linda’s front door; the characters are viewed in a medium long shot. Giving “night” as the first cue of the new scene seems to emphasize not only the physical condition of the new setting (darkness) and the change of setting (from interior to exterior) but also the temporal transition (from evening to night).

⁶ The first shot of Scene 75 of *Untergang* depicts an exterior long shot of civilians on a street, looking for something and then starting to run in all directions as a bomb alert begins to play. The previous scene takes place inside Hitler’s bunker, presumably at evening or night, so that the topicalized phrase *de día* (‘in the daytime’) functions in the same way as “night” in 11/AD-EN/*Dancer*.

The transition occurs when the shot of Hitler (9) walking in the bunker with his back to the camera is cut into a medium close-up of two officers also walking (10a), viewed from the side (another soldier stands in the background). The soundscape changes abruptly as Hitler's footsteps stop, and muffled, distant-sounding explosions and a buzz of conversation begin. The opening shot is continuous (illustrated by the dash line from 10a to 10b), meaning that the camera pans to follow the characters' movement, revealing more shot space, that is other people and parts of the setting. In 10b, which is the opening shot's final framing, the two officers merge into the crowd. This opening shot illustrates not only how a shot distance can focus on interactional space but also how spatial coverage can change during one shot (from the medium close-up in 10a to the medium long shot in 10b). Extracts 5 and 6 present the Spanish and German audio descriptions of the scene opening.

Extract 5: Scene 10/AD-ES/*Untergang*

(Scene 9)

HI: tonterías;
nonsense

((footsteps stop))

(Scene 10) [muffled explosions, buzz of conversation]

AD: **himmler y fegelein llegan**
Himmler and Fegelein arrive
a uno de los salones de la cancillería
at one of the halls of the Chancellery
lleno de otros altos mandos.
full of other high-ranking officers
((dialogue))

Extract 6: Scene 10/AD-DE/*Untergang*

(Scene 9)

HI: ach UNSINN;
ah, nonsense
((footsteps stop))

(Scene 10) [muffled explosions, buzz of conversation]

AD: **viele militärs in einem saal.** (-)
many officers/military people in a hall
tageslicht fällt herein. (-)
daylight enters

The AD-ES (Extract 5) begins with a reference to characters and their action, 'Himmler and Fegelein arrive', and then expands the view to the setting, 'at one of the halls of the Chancellery', ending with more information about the interactional space, 'full of other high-ranking officers'. The opening identifies the characters by name (Himmler, Fegelein) and the setting by a specific noun ((one of the halls) in the Chancellery). The first utterance about the new scene is thus a sentence in the active voice in which a reference to character action serves as the beginning point.

Contrasting the Spanish opening to the AD-DE version of the same instance (Extract 6) reveals interesting differences regarding referentiality. The AD-DE introduces the same scene by means of a focus construction: "viele militärs in einem saal" ('many officers/military people in a hall'). This focus construction indicates the salience of two aspects of space: 'many officers/military people' refers to characters, i.e. interactional space, and 'in a hall'

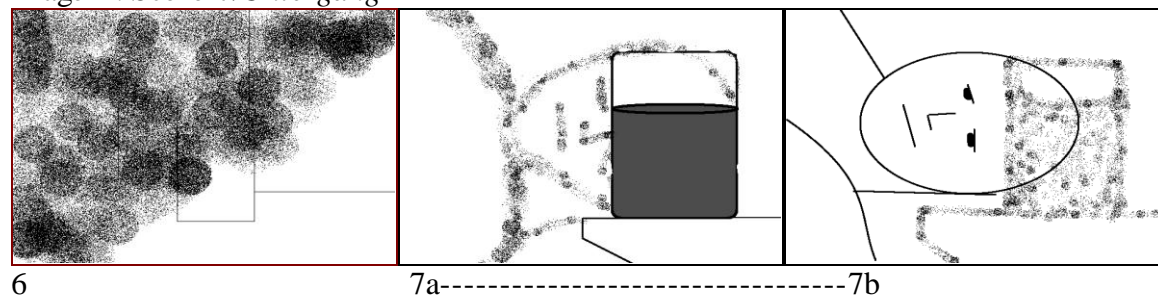
designates their location, i.e. setting. The two openings manifest difference in their approach to the referents they verbalize: the AD-DE describes the characters and the setting with a general category (military people; hall), with the modifier “viele” implying that many are seen but not specifying any one character, and excludes information about action. The Spanish opening offers more specific cues by identifying the characters and the setting. Structurally, it resembles the visual cueing in the original by referring first to the two characters and then to the setting and other people. Since neither the two officers nor the setting have so far been introduced in the film, the AD-ES deploys an explicative strategy by identifying the characters and the setting. That the German audio description excludes the more specific information at this point does not mean that it fails to give it completely; rather, the description may only be structured differently and more information may be rendered later (as in fact occurs). Here, the AD-DE opening anticipates the spatial information (the amount of people and the type of place) that is perceivable in the final framing of the opening shot (10b).

Interaction-driven AD openings typically focus first on interaction (on characters and their action) and, as such, can be seen as ‘interaction-driven’, but immediately afterwards verbalize more aspects of the shot space (e.g. the setting, as in Extracts 5 and 6). By doing so, they seem to suggest the listener expand the view of the scene from the character(s) to the environment. In contrast, the AD opening can be restricted to referring only to interaction (e.g. the focus construction “con los niños” (‘with the children’) in 130/AD-ES/*Untergang*⁷). Focus construction and topicalization are also used to introduce distinct spatial aspects at once; in the present study, such AD openings are termed ‘multiple-aspect AD opening’. This strategy occurs not only in interaction-driven scene openings (although it is most common here, cf. Table 4), and it is discussed with other deviant AD opening strategies in section 3.4 below.

3.3 Detail-driven scene openings

The last example discusses an opening that focuses on a detail in space. Image 4 illustrates a transition in *Untergang* with a detail-driven opening shot.

Image 4: Scene 7/*Untergang*



The exterior shot of bombs exploding (6) changes into a static, interior extreme close-up (7a) of a glass containing transparent liquid in the foreground and a blurred human figure in recumbent position in the background (the glass is on the right side of the frame with the blurry human figure in its background). The sounds of explosions that were loud in the

⁷ Scene 130 in *Untergang* continues the action initiated in a previous scene, in which Traudl meets Goebbels' (Hitler's minister) children and goes to look for food for them. Scene 130 begins by depicting Traudl in a medium long shot, spreading marmelade on a bread at a table, with the children seated around her. The Spanish AD cues the scene change with a short elliptical phrase “con los niños” (‘with the children’) (dialogue between Traudl and the children begins almost immediately in the new scene), which indicates the presence of the children as well as their familiarity with the definite pronoun “los”.

preceding scene become muffled in the new scene. This change implies a transition of some kind; here, the location changes from outside to inside. The liquid in the glass vibrates along with the sound. This image lasts about three seconds, after which the focus sharpens to the human figure (7b); we may now recognize the figure in close-up as a familiar character from a previous scene, Traudl. The view of space is restricted to a small object (the glass of water) during the first framing (7a) and widens to show the character's face and upper body at the second instance (7b). This restricted view is cued in the AD-DE (Extract 7) and in the AD-ES (Extract 8).

Extract 7. Scene 7/AD-DE/*Untergang*

(Scene 6) ((distant shouts + two explosions))

(Scene 7) [muffled explosions + weak gurgle]

AD: **auf einem nachtschränkchen, vibriert ein wasserglas.**
on a night table, a water glass is vibrating
 (-) licht brennt. (--)
 (-) (a) light is on
 traudl, (.) liegt im bett und öffnet die augen.
Traudl (.) is lying in bed and opens her eyes

Extract 8. Scene 7/AD-ES/*Untergang*

(Scene 6) ((distant shouts + two explosions))

(Scene 7) [muffled explosions + weak gurgle]

AD: **dentro del búnker,**
in(side) the bunker
el agua en un vaso vibra con las explosiones; (-)
the water in a glass is vibrating with the explosions
 traudl abre los ojos en el dormitorio.
Traudl opens her eyes in the bedroom

As in other openings with topicalization, in the AD-DE (Extract 7) the topicalized item 'on a night table' cues the change of scene as well as the saliency of a certain aspect, which in this case is an object in the setting, a night table. With the preposition "auf", it refers to the object as location: something is ON that table. The locative phrase is followed by a description of a physical event: 'a water glass is vibrating'. The whole utterance refers to single objects in space (night table and water glass), thus limiting the view to a detail in the scene. It does not verbalize the general setting or the characters that act in it; implicitly, however, 'night table' cues the setting as a bedroom since night tables are associated with bedrooms. In contrast, the Spanish version of the same opening (Extract 8) topicalizes a reference to setting and describes the detail afterwards: 'in(side) the bunker the water in a glass is vibrating with the explosions'. Nevertheless, while the bunker as the setting for the scene is specified, the actual room is not (the glass could be anywhere). In the AD-DE, the setting (the bedroom) is implied by the reference to a typical piece of furniture.⁸

In addition to the devices presented in Extracts 7 and 8, three other types of detail-driven AD opening are worth mentioning. One is a focus construction that merely refers to an object and thus limits our mental view of the scene: for example, "das photo eines soldaten" ('the photo of a soldier') (61/AD-DE/*Untergang*)⁹. The other is a topicalization that foregrounds a word

⁸ I wish to acknowledge the editors for pointing out this interesting aspect.

⁹ Scene 61/*Untergang* opens with the camera panning objects on a desk in a close-up, one of which is a photo of a soldier. In the voice-over that immediately follows, Goebbels' wife addresses her son in a letter: "mein

or construction describing a facial expression: for example, “beklommen nähert sich Traudl Hitler’s wohnraum” (‘anxious, Traudl approaches Hitler’s apartment’) (140/AD-DE/*Untergang*¹⁰). The third opening type points to a detail in a larger setting: for example, “Jeff’s truck trundles along a snow covered road” (61/AD-EN/*Dancer*¹¹). The last two types deal with spatial aspects that can be considered a detail also in interaction-driven or setting-driven openings: facial expressions and moving objects.

3.4 Deviant AD opening strategies

In addition to setting-, interaction- and detail-driven AD opening strategies, the material contains AD openings that deviate from these three basic strategies. Two of the deviant strategies share characteristics with the basic ones (discussed in 3.1-3.3 above), while one strategy is substantially distinct.

The first deviant strategy is the ‘multiple aspect strategy’ already introduced in section 3.2. It is characterized by the way it simultaneously foregrounds different spatial aspects: in the example given (Extract 6), interaction (‘many officers/military people’) and setting (‘in a hall’). Sometimes an opening description begins with a reference to time; it is thus a ‘time-driven AD opening’. This strategy introduces a new scene by an explicit indication of temporal transition that, however, may be followed by a spatial cue (e.g. “después Eva escribe a máquina” (‘later, Eva is writing at a typewriter’) in 56/AD-ES/*Untergang*¹²). The final deviant strategy is dramatically different from all the other strategies: ‘no AD opening’ occurs when no verbal description is offered at a scene transition. This may be due to prevailing dialogue (allowing no time or space to describe), and on some occasions, especially when the story returns to a familiar scene, the auditory narration may offer sufficient spatial cues (e.g. to recognize the characters in dialogue or a familiar setting).

Neither the deviant nor the basic AD opening strategies are bound to any particular opening shots, as will be seen in the quantitative analysis (cf. in particular Table 4). Whereas the comparative analysis described different ways of opening a scene and defined corresponding AD strategies, the following quantitative analysis studies the variety of representational strategies and linguistic devices that was found in the comparative analysis.

4. The quantitative analysis

The quantitative analysis has three parts. The first part discusses the frequency of different shot and AD opening strategies in the material, and the second part describes the distribution of AD opening strategies relative to the frequency of filmic opening strategies. The final part provides a summary of the central linguistic devices used in the AD opening strategies.

geliebter sohn...” (‘my beloved son...’). Thus, referring to the photo in the audio description links up with the content of the voice-over.

¹⁰ Here, the first shot of Scene 140 is a medium close-up of Traudl, framed frontally from a slightly low angle, walking slowly in Hitler’s office. The German audio description topicalizes an emotional state, “beklommen” (‘anxious’), which is typically observable on the face, and then enlarges the view to describe character action and setting.

¹¹ Scene 61 in *Dancer* moves the story from an interior setting in Scene 60 to an outside location, depicted in a long shot of a moving truck in a snowy environment. The AD-EN describes the new scene, first identifying the moving object in it as “Jeff’s truck trundles” and then referring to the environment with “along a snow covered road”.

¹² Scene 55 and Scene 56 in *Untergang* are consecutive but their relation is elliptical, in other words, the story “jumps” from one situation to another: in Scene 55, Eva, Hitler’s fiancée, is chatting with other people in Hitler’s apartment, and in the next scene, 56, she is sitting at her desk. Thus, the Spanish AD explicates the shift by the temporal adverb “después” (‘later’).

4.1 Frequency of shot and AD opening strategies

In total, 241 opening shots were analysed in the two films. Thus, 482 AD openings were analysed since this study worked with two different audio descriptions of the two films (German and Spanish in *Untergang*; English and German in *Dancer*).

Table 3 gives an overview of the data. The information in the table is briefly described here, and a more detailed explanation follows the table. The left column presents the types of opening strategy in six categories: ‘setting-driven’, ‘interaction-driven’, and ‘detail-driven’, all of which are shared by both film and AD, and ‘multiple aspect AD openings’, ‘time-driven AD openings’ as well as ‘no AD opening’, which are AD-specific opening strategies. The proportion (number and percentage) of each shot opening strategy is given in the middle column, and the corresponding incidence of AD opening strategies appears in the right column.

Table 3: Overview of AD openings

	Source text (film)	Target text (AD)
Scene openings in total	241 (100 %)	482 (100%)
Setting-driven openings	90 (37.3)	179 (37)
Interaction-driven openings	112 (46.5)	194 (40)
Detail-driven openings	39 (16.2)	27 (6)
Multiple aspect AD openings		10 (2)
Time-driven AD openings		21 (4)
No AD opening		51 (11)

Interestingly, the proportions of setting- and interaction-driven strategies in the source and target text are very similar: of the total amount of both opening shots and AD openings, 37 % are setting-driven; for interaction-driven openings, the scores are 46.5 % of opening shots and 40 % of AD openings. Thus on the basis of this data, scenes most often begin - both in the visual narration of the film as well as in AD - by focusing either on setting or interaction. The occurrence of detail-driven openings is rather different: 16.2 % of the opening shots are detail-driven, while only 6 % of the AD openings are classified as such. The other two AD strategies are significantly less frequent: the time-driven strategy occurs in 4 % and the multiple-aspect strategy in 2 % of the AD openings. The relatively high score of openings with no AD is likely to be caused by the high number of scenes beginning with dialogue, which impedes the use of AD. Other reasons explaining this distribution will be discussed in conjunction with the occurrence of the filmic strategies in the next section.

4.2 Frequency of AD opening strategies relative to filmic strategies

In order to analyse the correspondence of the AD opening strategies to the strategies in the source text, Table 4 shows the frequency of each AD opening strategy relative to the three shot opening types of filmic representation: setting-, interaction- and detail-driven openings.

Table 4: Frequency of AD opening strategies *per* shot type

Frequency of AD opening strategies	<i>per</i> setting-driven opening shots	<i>per</i> interaction-driven opening shots	<i>per</i> detail-driven opening shots
Setting-driven AD openings	74	79	26
Interaction-driven AD openings	68	102	24
Detail-driven AD openings	17	3	7
Multiple aspect AD openings	2	7	1
Time-driven AD openings	8	7	6
No AD opening	10	27	14

While some significant correspondence was found in the overall frequencies of filmic and AD strategies (recall the proportions of setting- and interaction-driven strategies in the source and target texts in Table 3), the comparison between AD strategies and shot types reveals that the AD and shot opening strategies do not always match. For instance, setting-driven AD openings are not exclusively used to introduce setting-driven scene openings. In fact, the setting-driven strategy is used to almost the same extent in both the films' setting- and interaction-driven openings, being slightly more frequent with interaction-driven openings (74 for setting-driven shots against 79 for interaction-driven shots). This means that, for example, an opening focusing on characters (e.g. through a medium long shot) is not always reflected in AD, where the first cue is often a reference to a wider aspect, the setting. Or to illustrate with another example: scenes beginning with a detail-driven shot (close-up or extreme close-up) are more often introduced with the setting- or interaction-driven strategy or with no audio description at all than with a detail-driven AD opening. A third observation involves detail-driven AD openings: most of them are located in setting-driven shot openings, and this may be due to a technical issue: the AD openings beginning with a reference to an object (e.g. a moving car) were classified as detail-driven.

The relatively high frequency of setting- and interaction-driven AD strategies implies that either the place of action or character(s) performing an action are often taken as anchoring points for the new scene. That place is the beginning point not only in setting-driven but also in interaction- and detail-driven openings may be explained by the need to cue a change (of place, but also of scene) explicitly (as predicted by Kluckhohn 2005), thus marking a transition with a locative expression. The fact that (inter)action is frequently salient in other openings than interaction-driven may have at least two explanations. In setting-driven openings, characters and their action may attract attention and thus become salient cues to AD, or the character is regarded as a narrative link between scenes and is used to create coherence, as when, for instance, the story returns to narrate action that was already happening in an earlier scene). In detail-driven openings, on the other hand, the focus on a detail may have an aesthetic rather than a narrative function and thus be considered less relevant than the character action that is framed after it (e.g. the opening shot is a close-up of a gun but the opening AD refers to the character loading the gun). This may in part explain why the detail-driven openings are less frequent in the AD openings (16 % of the opening shots but only 6 % of the AD openings are detail-driven, Table 3). The 'No AD' category is the most prevalent in interaction-driven openings, which, as I mentioned before, suggests that opening shots focusing on characters often include dialogue (cf. Straßner 2001:1095).

4.3 Variety of linguistic devices used in the AD opening strategies

Finally, it is worth noting that the AD opening strategies encompass a variety of distinct opening types. As the scope of this article does not provide enough space to present all types, I will cover only the types that relate to the setting-, interaction- and detail-driven strategies. They will be discussed next with examples from the data, most of which were already seen in the analysis. The types differ in the way they present spatial (or temporal) aspects as the first cue (setting, interaction, detail) and in their structural realization. The different structural devices are: focus construction, topicalization, active and passive voice sentence and 0-subject construction. The devices are generally not restricted to any one strategy (except for 0-subject construction, which in this data only appears with a setting-driven reference) but are deployed in different strategies; topicalization, for instance, is deployed in setting-, interaction-, detail-driven and multiple aspect openings. The devices are not language-dependent either although, in the present data, 0-subject constructions were found only in the German audio descriptions.

The setting-driven AD strategy encompasses opening descriptions that refer to the setting or to an element in it (such as illumination) as the first cue of the new scene. The different types are (in order of frequency):

- Reference to setting with focus construction, e.g. “im verwüsteten garten hinter der alten reichskanzlei” (‘in the devastated garden behind the old Reich Chancellery’)
- Setting (topicalization): “fuera del búnker los soldados se cuadran ante la salida de hitler” (‘outside the bunker the soldiers are standing to attention as Hitler exits (the bunker)’)
- Setting element (0-subject construction): “es dämmert” (‘it is getting light’)
- Setting element (topicalization): “in the gray light of dawn (.) selma walks steadily along the track”
- Setting (active voice): “der wald lichtet sich” (‘the forest thins out’)

The interaction-driven AD strategy opens the description of the scene by referring to character (inter)action or to a detail about this action:

- Interaction (active voice): “himmler y fegelein llegan a uno de los salones de la cancillería lleno de otros altos mandos” (‘Himmler and Fegelein arrive at one of the halls of the chancellery full of other high-ranking officers’)
- Interaction (topicalization): “holding her bike (.) selma waits for a train to pass at the crossing near to her trailer”
- Interaction (focus construction): “con los niños” (‘with the children’)

The detail-driven AD strategy foregrounds a detail, such as an object in the setting:

- Detail (active voice): “jeff’s truck trundles along a snow-covered road”
- Detail (passive voice): “the figures of the workers are silhouetted in the light of the doorway as they leave the factory”
- Detail (topicalization): “auf einem nachtschränkchen vibriert ein wasserglas” (‘on a night table a water glass vibrates’)
- Detail (focus construction): “das photo eines soldaten” (‘the photo of a soldier’)

These AD opening strategies resemble the filmic strategies of shot distance by cueing the salience of different spatial aspects (setting, interaction, or detail). The categorization is, however, not without problems. While some of the AD types contain a relatively evident and simple reference to *one* spatial aspect (e.g. ‘it is getting light’), other AD openings, in particular those in active or passive voice or with a topicalized item, entail a string of

references, each of which may cue distinct aspects of the scene (recall the AD-ES opening in Extract 5: ‘Himmler and Fegelein arrive...’). Accordingly, some of the opening types have been classified as the multiple-aspect AD strategy. The same problem occurs with some detail-driven opening types, where the challenge lies in defining a detail and its referential coverage in the scene. For example, in the aforementioned example of the AD type ‘detail (active voice)’, ‘jeff’s truck trundles along a snow-covered road’, the reference classified as detail, ‘Jeff’s truck’, reflects, in this context cueing both action (‘trundles’) and setting (‘along a snow-covered road’), more an element in the setting (a moving object) than a detail in the sense of close-up (e.g. some part of the truck). And indeed, this AD opening occurs in a setting-driven scene opening in *Dancer* (scene 61). Thus, as the first cue ‘Jeff’s truck trundles’ seems to direct our attention to an element in the setting, it could actually be classified as realizing the setting-driven AD strategy. Other AD openings referring to a detail seem to reflect a closer view and focus on, for instance, a face or a small object (recall the analysis of detail-driven openings). Thus, the attentional variation that the visual mode enables seems to create variation in the AD as well, which the discussion of Extracts 1 and 2 also illustrates. Moreover, other visual (and auditory) elements of filmic narration also have an influence on verbalization, which may in part explain the variation in the AD openings examined in this study (for example illumination, cf. Extracts 3 and 4).

5. Concluding remarks and discussion

This article aimed to describe strategies of spatial cueing in film and AD as well as defining links between the visual and verbal representation of space that could be beneficial to the translation of images into words. Looking at shot distance as a relevant spatial cue in the filmic source text, a comparative analysis was conducted in order to study features of this cue in the translations, i.e. different-language audio descriptions. The analysis demonstrated that similar strategies of spatial cueing may be put to use in the visual source and the verbal target texts and that language provides different devices for the cueing. This similarity seems to build partly on objective evidence (shot distance as relevant cue) and partly on the judgments by the audio describers (e.g. what is verbalized from a setting-driven shot) (cf. Chesterman 1996:159). Any shot can display multiple aspects to be verbalized in AD, but shot distance is one strong cue. The quantitative approach exhibited a variety of strategies and devices found in the AD material at hand and showed degrees of correspondence as well as dissimilarity between filmic and audio descriptive strategies of scene opening.

As a result of the intermodal comparative analysis, the following strategies and devices of filmic AD in spatial cueing could be defined:

- With the device of shot distance, which controls the amount of space viewed in the image, different aspects of shot space may become the point of interest (e.g. the setting or the action of characters). AD may compensate this variation by reflecting the *spatial coverage* in its *referential coverage*: for instance, if the shot frames a setting, it may be referred to as an entity (e.g. a garden).
- On the linguistic-structural level, a variety of constructions may be used. With topicalization and focus construction, any aspect of space may be made the beginning point for the new scene and the change (of scene) emphasized. Various aspects of space may be made topical, from the characters and their action to the setting, its elements and a detail.
- On the semantic-referential level, distinct spatial aspects can be referred to with more or less precision, for instance varying between proper names and common nouns. The implicitness of visual representation can sometimes be cued by ambiguous verbal references.

To sum up the quantitative analysis, we observed that the filmic representation and the audio descriptions converge in the proportion of two strategies in the overall sample, beginning most scenes with either setting- or interaction-driven strategy. Notwithstanding this similarity, the comparison between AD and filmic strategies revealed a considerable discrepancy in the use of both strategies. The comparison also evoked reflection about the possible motivations, narrative and aesthetic, on the use of different strategies. This relates to more general issues about what motivates AD, i.e. which aspects of the narrative, not merely spatial or visual, are considered relevant to be verbalized, and which possible consequences do these choices have with regard to, on the one hand, the mediation of the source text and, on the other, the reception of the target text (about source-text based motivations, see also Pérez Payá 2007). If one function of AD is to try to convey not only *what* is narrated but also *how* it is narrated, then choosing the narrative motivation over the aesthetic may produce contradictions.¹³ For instance, the order in which elements in the shot are described could reflect iconically the way in which they are made salient and in which the audio describer - as a viewer - observes the image. However, some relevant narrative cues, such as the change of scene, may be disregarded at the same time. Having said this, it is worth noting that the present research did not aim to evaluate motivations or the choices between them but rather to describe different AD strategies and devices and discuss their relationship of similarity to the filmic source text. Based on the samples analysed here, language does indeed supply AD with multiple means of representing the audiovisual source text.

While the study focused on one feature of the source text (shot distance), it recognizes the role of other audiovisual elements in producing AD. Therefore, further research that takes the meaning-making complexity of film into consideration is needed. A dynamic view is relevant since shots are interrelated and interpreted as narrative continua. Research on narrative cueing through sound would also be profitable (cf. Fryer 2010). Another rather unexplored dimension is the vocal delivery of AD (see e.g. Braun 2008:24); for instance, prosody might play a significant role in fore- and backgrounding (spatial) information.

The findings in this article shed light on both theoretical and practical issues of AD. On a general note, they point to relations of similarity in intermodal translation, as links between visual and verbal representation were found (strategic similarity in spatial cueing). Furthermore, they may benefit AD practice by providing knowledge of visual and verbal cueing on which AD strategies and devices may build; one may ask, for instance, whether certain linguistic structures could be recycled in filmic AD in the same systematic way as certain visual devices are deployed in films.

Acknowledgements

I wish to thank Professor Liisa Tiittula and Professor Henry Bacon as well as the two anonymous reviewers for the helpful and valuable comments on an earlier version of this paper.

Author's Address:
Department of Modern Languages
Translation of German
PL 24
00014 University of Helsinki
Finland

¹³ I am grateful to one of my anonymous reviewers for pointing out this issue.

Appendix: Transcription symbols

Adapted from the GAT(2) transcription system (Selting et al. 2009) and its English version by Couper-Kuhlen & Barth-Weingarten (2011).

FE:	Name of speaker (abbreviated)	AD (audio describer), FE (Fegelein), KA (Kathy), HI (Hitler), SE (Selma), UM (unknown man)
[]	Speech or sound that overlaps with the following utterance	
((suspire))	Sound or non-verbal vocal action	
UNSINN	Capital letters illustrate word or syllable stress	
(.)	Short silence or micro pause (estimated <0.5 sec)	
(-) (--)	Longer silence or short pause (estimated >0.5 sec)	
	Final pitch movements:	
?	Rising to high	
,	Rising to mid	
-	Level	
;	Falling to mid	
.	Falling to low	

Filmography

Dancer in the Dark (2000). Directed by Lars von Trier. Zentropa Entertainments. [Audio described in English by IADA Ltd., 2000; audio described in German by DBSV / Projekt Hörfilm, 2001]

Der Untergang / El hundimiento (2004). Directed by Oliver Hirschbiegel. Constantin Film. [Audio described in German by Bayerischer Rundfunk, 2005; audio described in Spanish by ONCE, 2006]

References

- Benecke, Bernd (2004) 'Audio-Description', *META*, 49(1): 78-80.
- Benecke, Bernd (2007) 'Audio Description: Phenomena of Information Sequencing', in Heidrun Gerzymisch-Arbogast and Gerhard Budin (eds) *MuTra 2007 – LSP Translation Scenarios. Conference Proceedings*, Vienna, 2008. Available online at [http://www.euroconferences.info/proceedings/2007_Proceedings/2007_proceedings.html] (accessed 17 March 2011).
- Blakemore, Diane (1992) *Understanding Utterances*, Oxford: Blackwell Publishing Ltd.
- Bordwell, David (1985) *Narration in the Fiction Film*, Madison: University of Wisconsin Press.
- Bordwell, David and Kristin Thompson (1990) *Film Art: An Introduction*, New York: McGraw-Hill.

- Bourne, Julian and Catalina Jiménez (2007) 'From the visual to the verbal in two languages: a contrastive analysis of the audio description of *The Hours* in English and Spanish' in Jorge Diaz-Cintas *et al.* (eds), 175-187.
- Braun, Sabine (2007) 'Audio Description from a discourse perspective: a socially relevant framework for research and training', *Linguistica Antverpiensia*, 6: 357-369.
- Braun, Sabine (2008) 'Audio Description research: state of the art and beyond', *Translation Studies in the New Millennium*, 6: 14-30.
- Brown, Gillian and George Yule (1983) *Discourse Analysis*, Cambridge: Cambridge University Press.
- Chesterman, Andrew (1996) 'On similarity', *Target*, 8(1): 159-164.
- Chesterman, Andrew (1997) *Memes of Translation*, Amsterdam/Philadelphia: John Benjamins.
- Couper-Kuhlen, Elizabeth and Dagmar Barth-Weingarten (2011) 'A system for transcribing talk-in-interaction: GAT 2. English translation and adaptation from M. Selting *et al.*: Gesprächsanalytisches Transkriptionssystem 2', *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 12: 1-51.
- Diaz-Cintas, Jorge, Pilar Orero and Aline Remael (eds) (2007) *Media for All: Subtitling for the Deaf, Audio Description and Sign Language*, Amsterdam: Rodopi.
- Fest, Joachim and Bernd Eichinger (2005) *Der Untergang. Das Filmbuch* [Edited by Michael Töteberg], Hamburg: Rowohlt Taschenbuch Verlag.
- Fix, Ulla (ed) (2005) *Hörfilm - Bildkompensation durch Sprache*, Berlin: Erich Schmidt Verlag.
- Fryer, Louise (2010) 'Audio description as audio drama: a practitioner's point of view', *Perspectives: Studies in Translatology*, 18(3): 205-213.
- Haspelmath, Martin (2008) 'Frequency vs. iconicity in explaining grammatical asymmetries', *Cognitive Linguistics*, 19(1): 1-33.
- Hirvonen, Maija and Liisa Tiittula (in press) 'Verfahren der Hörbarmachung von Raum: Analyse einer Hörfilmsequenz', in Heiko Hausendorf, Lorenze Mondada and Reinhold Schmitt (eds) *Raum als interaktive Ressource*, Tübingen: Gunter Narr Verlag.
- Hyks, Veronica (2005) 'Audio Description and Translation: Two related but different skills', *Translating Today*, 4: 6-8.
- Katz, Steven D. (1991) *Film Directing Shot by Shot: Visualizing from concept to screen*, Studio City, CA: Michael Wiese Productions.
- Kluckhohn, Kim (2005) 'Informationsstrukturierung als Kompensationsstrategie – Audiodeskription und Syntax', in Ulla Fix (ed), 49-66.
- Kress (2009/2011) 'What is Mode?' in Carey Jewitt (ed) *The Routledge Handbook of Multimodal Analysis*, 2nd. ed., London and New York: Routledge, 54-67.
- Kruger, Jan-Louis and Pilar Orero (eds) (2010) *Perspectives: Studies in Translatology. Special issue: Audio description, audio narration, a new era in AVT*.
- Lambrecht, Knud (1994) *Information Structure and Sentence Form: Topic, focus, and the representation of mental referents in discourse*, Cambridge: Cambridge University Press.
- Levin, Daniel T. and Caryn Wang (2009) 'Spatial representation in cognitive science and film', *Projections*, 3(1): 24-52.
- Matamala, Anna and Naila Rami (2009) 'Análisis comparativo de la audiodescripción española y alemana de *Good-bye Lenin*', *Hermeneus*, 11: 249-266.
- Monaco, James (2009) *How to Read a Film*. Oxford and New York: Oxford University Press.
- Morgner, Henrike and Steffen Pappert (2005) 'Darstellung des Untersuchungsmaterials', in Ulla Fix (ed), 13-32.
- Nord, Christiane (1991) *Text Analysis in Translation*, Amsterdam: Rodopi.

- Orero, Pilar (2007) 'Sampling audio description in Europe', in Jorge Diaz-Cintas *et al.* (eds), 111-125.
- Pérez Payá, María (2007) 'La audiodescripción: traduciendo el lenguaje de las cámaras', in Catalina Jiménez Hurtado (ed) *Traducción y accesibilidad*, Frankfurt am Main: Peter Lang, 81-91.
- Remael, Aline and Gert Vercauteren (2007) 'Audio describing the exposition phase of films. Teaching students what to choose', *Trans - Revista de traductología*, 11: 73-93.
- Smith, Michael B. (2002) 'The polysemy of German *es*, iconicity, and the notion of conceptual distance', *Cognitive Linguistics*, 13(1): 67-112.
- Seiffert, Anja (2005) 'Räumliches Hören. Eine schemaorientierte Analyse der audiodeskriptiven Darstellung der Handlungsräume', in Ulla Fix (ed), 67-86.
- Selting, Margret, Peter Auer, Dagmar Barth-Weingarten, Jörg Bergmann, Pia Bergmann, Karin Birkner, Elizabeth Couper-Kuhlen, Arnulf Deppermann, Peter Gilles, Susanne Günthner, Martin Hartung, Friederike Kern, Christine Mertzluft, Christian Meyer, Miriam Morek, Frank Oberzaucher, Jörg Peters, Uta Quasthoff, Wilfried Schütte, Anja Stukenbrock, Susanne Uhmann (2009) 'Gesprächsanalytisches Transkriptionssystem 2 (GAT2)', *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 10: 353-402 (www.gespraechsforschung-ozs.de).
- Snyder, Joel (2005) 'Audio description. The visual made verbal across arts disciplines - across the globe', *Translating Today*, 4: 15-17.
- Straßner, Erich (2001) 'Kommunikative und ästhetische Leistungen der Sprache in der Geschichte des Films', in Joachim-Felix Leonhard, Hans-Werner Ludwig, Dietrich Schwarze, Erich Straßner (eds) *Ein Handbuch zur Entwicklung der Medien und Kommunikationsformen, Teilband 2*, Berlin and New York: Walter de Gruyter, 1093-1106.
- Stöckl, Hartmut (2004) *Die Sprache im Bild - Das Bild in der Sprache*, Berlin: Walter de Gruyter.
- Trier, Lars von (2000) *Dancer in the Dark*, London: Film Four Books.