Imaginary places –
Verbalization of setting in the English, German and Spanish audio descriptions of *Slumdog Millionaire*

Maija Hirvonen

University of Helsinki

Paula Igareda

Universitat Autònoma de Barcelona

Abstract

In this article we want to demonstrate how three different audio descriptions of the same film construct setting, i.e. the environment or place where the story happens. Audio description is an intermodal translation in which visual information is verbalized and uttered in spoken form. It helps visually impaired people to better access visual and audiovisual communication and culture. The study presented here describes the use of language as verbal cues for setting the film's scene in the mind of the receiver; for example, the same setting may be verbalized as "interrogation room" or "a small cranked room". Different versions of the same filmic event provide insight into the possibilities and limits of audio description in verbalizing visual experience by exhibiting differences and similarities in the strategies of verbalization. For this aim, we carried out a comparative analysis of three different-language audio descriptions (English, German and Spanish) of the film *Slumdog Millionaire* (Danny Boyle & Loveleen Tandan 2008). We will illustrate this analysis in the article by discussing the gradual construction of setting in one scene.

Keywords: audio description, film setting, verbalization, Slumdog Millionaire

1. Introduction

Describing, or the verbal account of an extra-linguistic experience, is an interesting area of language studies. On one hand, it is seen as a more or less uninhibited communicative act that, for instance, serves to answer a question ˙What happened?', however the description may be steered by focussing questions (see e.g. von Stutterheim & Kohlmann 2001 and Carroll & von Stutterheim 2007). Describing can be studied as naturally occurring interaction (see von Stutterheim & Kohlmann 2001) as well as used as a method in cross-linguistic experiments, in which case the differences and similarities between languages are particularly intriguing (Carroll & von Stutterheim ibid., Orero 2008). A further approach to describing is to apply

it as a form of translation. This is the focus taken in the present article, and the type of describing that is studied is audio description.

Audio description (AD) means verbalizing visual information. It is intermodal translation that intends to overcome a sensory boundary that visually impaired people encounter with regard to visual and audiovisual communication and culture (e.g. theatre, movies, and visual art). AD can be deployed in a variety of situations and contexts, ranging from 'live' performances (e.g. describing a theatre play or a museum exhibition during a guided tour) to pre-recorded texts (e.g. describing a film for television or DVD) (Orero 2005). In the present article, we focus on the latter form, AD of films (termed henceforth 'filmic AD'). In filmic AD, the verbal description is constrained by the filmic soundtrack which sets out gaps (e.g. in the dialogue) during which descriptions can be uttered. As in any verbalisation, in AD film's (audio-)visual stimuli are conceptualised and categorised with linguistic means; concrete representations are turned into abstractions (cf. Schwarz 1992: 104–105). When describing, one must conceptualize, i.e. interpret the state of affairs, and verbalize that conceptualization (Stutterheim & Klein 2002). In the same way, receivers of AD make use of concepts, and linguistic expressions mobilize concepts in the mind (cf. Cruse 2011: 27). As Cruse puts it, "the essence of a concept is that it enables one to treat a range of objectively different experiences as being instances (or 'tokens') of one and the same type of experience" (Cruse 2011: 53). The linguistic level and the concepts it triggers are thus relevant in AD in that they serve as intersubjective links between communication partners (cf. Braun 2008).

In this article, we study verbalization in three different-language audio descriptions of one narrative film, *Slumdog Millionaire* (Danny Boyle & Loveleen Tandan 2008). More particularly, we are interested in how the setting – the physical surroundings and the place of action in the story – is conceptualized. We approach the conceptualization by the analysis of words and concepts: Do different languages (German, English, and Spanish) render different names for setting and in what aspect do these differ, or can we perhaps find similarities? The aim is to describe and define strategies of verbalization in AD and to discuss how AD functions as a trigger for a recipient's imagination and, thus, cue an imaginary space (see Hirvonen & Tiittula 2012). In this article, we approach the multilingual data (the three audio descriptions) by analysing the gradual construction of setting in one scene.

This article is organised into the following parts: Firstly, it sets up the theoretical framework regarding narrative films and the setting, in relation to AD. Secondly, it

introduces the research questions and aims of the study along with some background studies, the methodology and the corpus of study. Thirdly, it presents the analysis, illustrating it with one example scene. Finally, the conclusions are offered, together with the findings obtained and a critical evaluation of the study and future proposals of research.

2. Setting the scene

In narrative films, a basic entity of narration is the setting, which locates the story in a place and functions as scenery for action (see e.g. Bordwell & Thompson 1990). Hence, setting relates to the physical and architectonic notion of space which, in contrast to interactional space, exists *per se* but can also be the stage for human interaction (Schmitt & Deppermann 2010: 227): for example, a room or a city. These environments are socially codified in the sense that certain activities are expected to happen in them (ibid. 228): for example, chairs in a room imply sitting. In film, the construction of setting is often based on visual cues in the sense that the cues in *mise-en-scène* (staging) enable the identification of setting as a certain (type of) place (e.g. hospital gear in a room may imply that the setting is a hospital) and trigger expectations about action. In addition, the cues in cinematography and editing regulate the structure and amount of the setting that is visible in a shot. Furthermore, film's auditory narration – the acoustic properties of sounds and dialogue, for instance – cues space as well: for example, the volume and timbre of sound may suggest physical dimensions of space (e.g. an echo is produced in large places). For the sighted spectator, the visual cues of setting are concrete representations that must be verbalized in AD in order to be delivered in speech. On the other hand, AD recipients deploy the verbalizations to evoke ideas in mind as well as have direct access to the auditory cues that potentially support (but sometimes also contradict) the verbal cueing (see Fryer 2010). Therefore, not only the verbalization that describes setting with words but also the soundtrack must be taken into account in an analysis of setting representation in AD. On the other hand, the availability of the interpretations based on sounds is challenged, if fabricated sound effects are used, of which the audience does not have memories or experience (Remael 2012).

The idea of setting entails manifold aspects. It is not only constituted from the general constellation (e.g. a room) but also of its (typical) items (e.g. a chair, a table, etc.). Moreover, the constellation of setting is stratified and has different layers: it

can be a city that includes buildings, which have rooms, etc. The constituents have qualities (e.g. dimensions, colours, lightning) and typologies (a room can be a kitchen, an office, a hall, etc.). Setting also includes a social-interactive dimension, the functionality of a place (e.g. what action it typically involves). Hence, setting can be seen as triggering both physical and social information of the narrative.

With a film as source text, AD translates the visual, and occasionally auditory, information into language. The verbalization is then rendered in speech along with the film's original sound track, so that the sound track remains as undisturbed as possible. With regard to verbalization, central questions in filmic AD are what information should be verbalized explicitly (and what, consequently, can be understood from context or on the basis of sound track) and what verbal expressions to use (cf. Vercauteren 2007).

The basic difference between watching a film audiovisually, and listening to an audio described, i.e. audio-verbal-vocal version, can be illustrated with Figure 1:
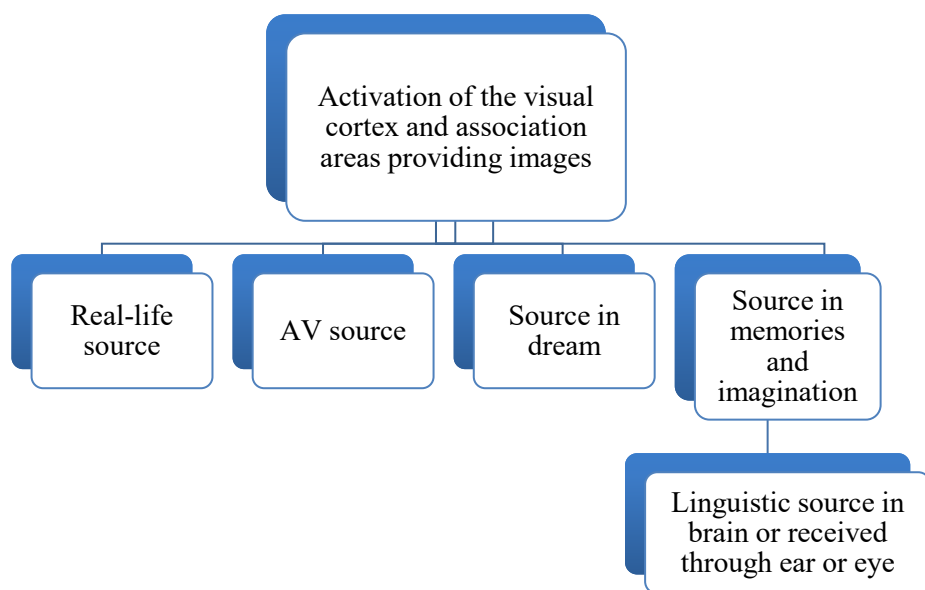

Figure 1. Grodal (1997: 23)

According to Grodal, "part of the mental activation is the same whether a given visual experience is caused by a ʿrealʾ experience, or by a memory of an experience, possibly activated by means of language, or caused by seeing a film or other audiovisual sequence" (Grodal 1997: 23). Thus, in an audio described film, the

mental representation of a visual event is activated by a linguistic source and received through speech, but it is also cued acoustically by the sound effects that are understood as coming from the events in the film. While film as an audiovisual experience directly triggers a visual representation in the mind, acquiring this with the linguistic resource necessitates a re-routing through the conceptual depository, that is, through sources in memories and imagination. Thus, verbal cues function as inferences for the re-construction of a filmic event in the mind.

3. The study

3.1 Research questions and aims

Inspired by the link between words and mental representations, we want to study filmic AD as verbalization, and to discuss strategies in constructing the setting. We ask:

- How is the film's setting verbalized in AD?

Regarding this question, we are interested in the means of AD in verbalizing different aspects of setting so as to cue a representation in the mind of the recipient (cf. Braun 2007, Seiffert 2005). In the light of a multilingual corpus, an interesting additional question is:

- How is the setting verbalized in different AD versions (languages) of the same source text, and do these languages/versions reflect shared conceptualizations?

A third and necessary question stems from the audiovisual context of film, in which the verbal cues of AD do not function alone but form a coherent whole with the film's soundtrack and dialogue. Consequently, we also inquire:

- How does the auditory environment of film (sound, speech and music) attend to the verbalization?

The aim of the study is to describe strategies for the verbalization of setting, as well as discuss similarities and differences in different-language audio descriptions. The general aim is to find general strategies of verbalization rather than to highlight

differences in languages. This means that the results may be applicable to the description of other visual aspects and to AD in other contexts as well.

3.2. Background studies

Hirvonen and Tiittula (2012) analyze auditory modes (spoken language, sound, and music) in the making of spatial meanings in a German audio described film. This study found that the modes provide a variety of resources to construct an imaginary space of narration, which is the foundation for a non-visual, yet multimodal construction of space in an audio described film (the so called ʿimaginary spaceʾ, see Schmitt & Deppermann 2010). In the framework of cognitively oriented film theory (see e.g. Bordwell 1985, Bordwell & Thompson 1990), this idea fits well to the notion of 'narrative cueing', which describes the function of visual and auditory representations in film as cues on the basis of which spectators construct the story in their minds. Thus, we study an audio described film as a stand-alone product, concentrating on the imaginary space it offers.

As already mentioned, the auditory modes at stake are multiple. Like Seiffert (2005), we analyze the audio description of setting from the point of view of verbalization: what kinds of keywords are used to refer to the place. Fryer (2010) draws attention to the functions and sound effects as auditory cues; this is taken into account in our study as well. The effects of music are numerous as well as important for the reception of film (see e.g. Schmidt-Banse 2002). However, as non-experts, we decided not to embark very profoundly on the analysis of musical cueing (some findings about music in audio described film are presented in Hirvonen & Tiittula 2012).

To widen the perspective from monolingual data, this paper focuses on three parallel texts of filmic AD (i.e. different-language audio descriptions of the same filmic content). Similar analyses have previously been carried out by Bourne and Jiménez (2007) for English and Spanish, Matamala and Rami (2009) for Spanish and German, and Hirvonen (e.g. 2013) for English, German, and Spanish. While Bourne and Jiménez focus on the comparison of linguistic units, mainly lexical and syntactic constructions, Matamala and Rami are predominantly concerned with the translation of cultural references, showing among other things that certain strategies of interlingual translation may be in use in AD as well. Hirvonen uses two scenes from the same data (*Slumdog Millionaire*) to contrast the visual representation to the linguistic one.

3.3 Methodology and corpus

Our approach in this study is reception-oriented and comparative. The audio-described film was studied as a stand-alone product and as an offer to the recipient to construct an imaginary space on the basis of AD and soundtrack (see Hirvonen & Tiittula 2012). The comparative approach means analyzing different productions of the same referent (different-language audio descriptions of one film) in relation to each other. The *tertium comparationis* is the – in this study, invisible – extra-linguistic substance (see e.g. Chesterman 1998), i.e. the visual information in the film. Thus, we assume that the different audio descriptions mean to refer to the same extra-linguistic experience, which is the film, but not necessarily to exactly the same referents because a selection from the filmic imagery must be made.

The corpus consisted of three audio-described versions of the film *Slumdog Millionaire* (Danny Boyle & Loveleen Tandan, UK, 2008). The three audio descriptions are: English (produced by Pathé Distribution in London, 2009), German (Hörfilm GmbH in Berlin, 2009), and Spanish (Navarra de Cine, S.L. in Aranguren, 2009)[1]. The duration of the film is approximately two hours.

Slumdog Millionaire narrates the story of a young boy from the slums of Mumbai who manages to win the million Rupee prize in the TV game show ʻWho wants to be a millionaire?', as well as, after severe difficulties, come back together with the love of his life, a girl from the slums. It is one of the few films that has been audio described in more than one language. One possible explanation for this popularity is that it won the Oscar for the best feature film (plus, it was nominated in seven other categories) in 2009. Most of the film's dialogues are in English, with some in Hindi. This additional language is then either dubbed or subtitled. In the German and Spanish versions, the Hindi dialogues are dubbed, but in the English version it is a different speaker than that of AD that voices the subtitles. Interestingly, the story takes place in India, which to a large extend is culturally distant to all three language-cultures that have audio described the film (British English, Peninsular Spanish, and Standard German). Cultural distance may produce additional constraints to AD, producing a generalisation of culturally distant or unfamiliar objects in the verbalisation (Matamala & Rami 2009). Of the three language-cultures, the context of the film seems to be the closest to the British English version

---

[1] It should be noted here that the different audio descriptions are not interlingual translations of one another but intermodal translations of the film. We would also like to take this opportunity to thank Lee Bye from Pathé Productions and Martina Wiemers from Hörfilm GmbH for providing the original AD scripts for the purpose of this study.

due to the common history with India. However, another interesting and culturally bounding element is the TV show 'Who wants to be a millionaire?'. It is famous worldwide and has been imported by a large number of countries (including in the three language-cultures in focus here)[2].

In order to do research on dynamic, audio-verbal data, we first had to create a textual record of the three audio described films. We created a separate document to each version based on the original AD scripts and added dialogues and descriptions of sound effects and music. Once the documents were ready, we structured the texts into scenes to help the comparison. This resulted in 76 scenes in each version. We wish to emphasise here that a division of film scenes is necessarily a product of interpretation (see Bordwell & Thompson 1990) and serves merely the function of organization. We wanted to keep the scenes as clear as possible, comprising of narrative continua of temporally linear entities with one core activity. However, this was not an easy task, since the film's discourse plays with time on many levels. The story occurs on three levels of time: the past (Jamal's childhood), the story's yesterday (Jamal in the contest) and the story's today (Jamal at the police station). These three levels interchange and merge; for example, the past is sometimes narrated as present and other times in the form of a flashback. Therefore, the setting sometimes changes within the same scene due to a dream or flashback sequence, or because one core activity is continuously presented in different places (e.g. a character wandering the streets and inside of buildings).

After organising the texts into scenes, we tracked and coded verbal expressions of setting and sonic cues that imply it. The verbal expressions varied from words or short prepositional phrases (e.g. *im Studio* [in the studio], scene 4) to complete sentences (e.g. *The city sprawls with high-rise apartments and green parks juxtaposed with huge slums and shanty towns*, scene 43). We found both direct references (or 'overt actants', Chesterman 1998: 78), such as *in the studio*, and indirect or implicit references (or 'implicit actants', ibid.) that describe an action or object but simultaneously lead to inferences about setting, such as *on a runway* (scene 8) which implies a wider setting, in that "runway" is a characteristic part of "airport". The sonic cues include sound effects such as street noises. Film dialogues contained very little references on the currently visible setting (spatial deixis such

---

[2] See the Wikipedia entry about the international franchise: http://en.wikipedia.org/wiki/Who_Wants _to_Be_a_Millionaire%3F.

as *here* was sometimes uttered); rather, they seemed to make reference to an off-screen space, i.e. to places that were outside the situation at hand (e.g. Jamal and the show's host talking about Jamal's workplace, scene 5).

We illustrate the analysis in the next section. It is a detailed analysis of a setting construction in scene 63. As it was mentioned before, the corpus was divided into 76 scenes and all of them were reviewed. Within these 76 scenes we focused on seven that we found more remarkable for the aim of the study. Scene 63 was selected to illustrate this analysis because it comprises all the research questions exposed from the beginning, this would help the reader to better understand the characteristics and methodology of this study.

4. Analysis: Complex construction of setting

We want to study the complex construction of setting by analysing the audio descriptions of a whole scene. In film, the possibilities to audio describe vary according to the amount of significant sounds, and for the present analysis we have selected a scene which is relatively long and has abundant AD, thus merely a few lines of dialogue. It is scene 63 that occurs towards the end of the film.

The structure of the analysis is as follows: We first describe the acoustic cues and then analyse the audio descriptions chronologically, that is, as a sequence of utterances, in order to reflect the gradual construction of setting. In the utterances, verbal cues of setting (mostly nouns) are rendered in bold font so as to direct the reader's attention to relevant parts. At the end of this section, we apply schematic representations (images) to illustrate the gradual construction of setting and the conceptual scheme the descriptions trigger.

In order to refer to the different audio descriptions, we use the following coding: AD-ES means the Spanish AD, AD-EN the English AD, and AD-DE the German AD. Non-English descriptions are translated into English and this translation is rendered in brackets ([ ]). We codify setting as a conceptual constellation with capital letters (e.g. TRAIN refers to "train" as setting) and setting items with small letters (e.g. platform refers to STATION setting in the sense that it constitutes a typical part of it).

Acoustic cues of this setting are shared by all three versions. The very first cue of a new setting (relative to previous one, which is the game show) is a sonic one: honking of a train. The sound begins at the end of Prem's (PR), the show's host's utterance in the previous scene, thus constituting a sound bridge to the new scene

(see Monaco 2009). Sonic cues increase as the new scene begins: there is honking, humming and a buzz of distant talking. These sounds continue in the background of the AD utterances and constitute a sonic ambience for the scene (see Fryer 2010).

Verbal cues of the setting are analysed next. We present the audio descriptions of the scene in chunks, determined by longer pauses in AD, as well as juxtaposed in a table, that should facilitate their comparison.

	AD-ES	AD-EN	AD-DE
1	((WWM excitement sound)) PR: Los sueños de tantos... ((rustle of paper)) ((honking in background)) por los suelos. [Dreams of so many… on the floor]	((WWM excitement sound)) PR: Dreams of so many ((rustle of paper)) on the floor. ((honking in background)) Jamal looks out of a **train window** as **it** races through the **city**.	((WWM excitement sound)) PR: Der Traum so vieler Menschen ((rustle of paper)) in tausend Stücken ((honking in background)) [The dream of so many people in thousand pieces] Er wirft die Schnipsel hoch in die Luft. [He throws the snippet up in the air]
2	((humming, squeaking, buzz of talking in background)) La historia vuelve al pasado. [The story goes back to the past] (2.0)	((humming, squeaking, buzz of talking in background)) (8.0)	((squeaking, humming in background)) In der **Stadt**. [In the city/downtown]
3			Ein **Zug** rast über eine **Eisenbahnbrücke**. [A train speeds along a railway bridge]
4			((squeaking, humming, buzz of talking in background)) Jamal lehnt sich aus dem **Fenster des Zugs**. [Jamal leans out of the train window]

Table 1: Transcriptions of the different AD versions

The first cue of a new scene in AD-ES (line 2 in the table) is metatextual and introduces a change in the story by referring to the past. The new setting remains

implicit, inferable from the sound effects, whereas AD-EN and AD-DE provide an explication about the new scene by the first utterances (line 2). In AD-EN, TRAIN is the place of action for the Jamal character, and CITY the place of action for the racing train. The definite article, *the* allows to identify the reference, based on previous discourse, as the city of Mumbai. AD-DE begins the scene with an overt reference to this setting (*in the city/downtown*). This type of elliptical phrase can be used in AD to mark a change and to quickly designate a new scene (see Kluckhohn 2005). The first concept of setting is CITY. Then, the situation becomes more complex by a description of action that entails two elements of setting: train and railway bridge (line 3). Next, the "train" element converts into a setting as it designates TRAIN as the place in which Jamal leans out of the window (line 4).

Comparing the three versions, we note the following differences and similarities in setting construction:

- AD-ES excludes the concrete information about the new setting that is provided by both AD-EN and AD-DE.
- AD-DE explicates several elements of CITY, while AD-EN mentions only the train.
- TRAIN is both an element of CITY and a setting of its own in AD-EN and AD-DE. The linguistic changing between Figure and Ground functions, that is, between the role of acting subject or item (Figure) and the background for action (Ground), is also recognized in Schubert (2009: 63) and, with regard to AD, in Hirvonen (2013).
- AD-DE also explicates other elements of CITY such as houses.

The verbal references fix conceptualisations for the acoustic cues audible in the scene, so that the honking, humming and buzz of talking can be identified as features of TRAIN, STATION and CITY. After the first cue, the setting construction proceeds as follows:

5	Jamal llega en un **tren** a la **estación de Chhatrapati Shivaji**. (2.0) [Jamal arrives in a train in the station of Chhatrapati Shivaji]		**Häuser** gleiten rasch an ihm vorüber. [Houses slip rapidly past him]

6		Der **Zug** fährt in einen **großen Bahnhof** ein. [The train arrives in a large station]
7		Jamal steht an der **Waggontür** und springt auf den **Bahnsteig.** (2.0) [Jamal stands at the wagon door and jumps onto the platform]
8	**Es un impresionante edificio centenario de estilo gótico** que fue declarado patrimonio de la humanidad por la UNESCO y que hoy es **la estación con más tránsito de trenes de toda la India**. (1.0) [It is an impressive hundred-years-old Gothic building that was declared World Heritage site by UNESCO and that today is the station with most traffic in India]	The **train** pulls into a **large station** and he jumps off. (5.0)	
9	**Los andenes están saturados de gente en todas direcciones**. (4.0) [The platforms are overfull by people in every direction]		**Hunderte Menschen** drängen aneinander vorbei. (3.0) [Hundreds of people crowd past each other]
10	**Uno de los relojes marca las cinco y cuarto de la tarde**. (2.0) [One of the clocks shows quarter past five in the evening]	**Hundreds of passengers** alight from the **train** and throng the **platforms.**	Vor dem **Bahnhof.** [In front of the station]
11		((car, honking, humming in background)) **An enormous station clock shows five minutes to five.** (3.0)	((humming, car in background)) Das **prächtige Gebäude im Kolonialstil ragt hoch in den Himmel.**

			[The magnificent colonial-style building towers high up into the sky]
12			((honking, humming, buzz of talking in background)) **Eine Uhr auf dem Bahnsteig zeigt Viertel nach fünf.** (3.0) [A clock over the platform shows quarter past five]

Table 2: Transcriptions of the different AD versions (part 2)

Now, the scene becomes more concrete in the AD-ES as well (line 5). Two ideas of setting are cued: TRAIN and STATION. The station is identified with a proper name, Chhatrapati Shivaji, thus constituting a specific reference to a certain place. The following utterance (line 8) describes STATION, referring to its physical qualities (impressive, hundred-years-old, Gothic building) as well as functionality (station with the most train traffic in India). That the station is busy is confirmed by the following utterance (line 9) and inferable from the sound of conversations as well. The general category *gente* [people] can be understood as referring to human beings as ʿadd-onsʾ that serve as decoration (cf. Monaco 2009) and as a typical element of STATION. The next utterance (line 10) also refers to a typical item in STATION, namely clock/clocks.

In the AD-EN, the TRAIN setting turns into an acting subject, after which a third setting, STATION, is introduced (line 8); this syntactic repositioning changes the Figure and Ground relations in the space (see Hirvonen 2013). All in all, the STATION setting becomes more complete with references to typical elements: passengers, train, platforms, and clock. The size of the setting is also indicated, first explicitly by the adjective *large* and then implicitly by the descriptions *hundreds of passengers* and *enormous station clock*.

In the AD-DE (line 5), the first utterance of this part refers to CITY again in that houses are understood as a typical element of cities. Then, the train becomes an element again, and a new setting, STATION, is introduced (line 6). The following description refers to character action, bearing references to the setting (line 7): WAGON DOOR and PLATFORM. The next utterance describes STATION similarly to AD-EN, with "people" as decoration (line 9). Then, a change in place is marked again by a focus construction (line 10): *in front of the train station*. Interestingly, such information is not provided in either of the other two AD

versions. This description invites to view the station from the outside, thus taking us back to CITY setting in which the station is one typical element. This element is described in a separate utterance (line 11), and the description resembles the corresponding AD-ES description in that it refers to physical qualities (size and appearances): *prächtig*, *im Kolonialstil, hoch*. Even a metaphoric expression is used to describe the size: *(hoch) in den Himmel ragen*. As both AD-ES and AD-EN do, AD-DE also designates the "clock" element as a separate description (line 12).

The above analysis studied the construction of setting until the moment in which all three versions describe the clock showing a particular time of day. After this, the focus in all audio descriptions shifts (back) to Jamal's action.

The setting of the scene 63 cued through the audio descriptions is illustrated in the Figure 2 images.

La historia vuelve al pasado reciente.
[the story goes back to the recent past]

Jamal llega en un tren a la estación de Chhatrapati Shivaji
[Jamal arrives in a train in the station of Chhatrapati Shivaji]

Es un impresionante edificio centenario de estilo gótico...
[(it) is an impressive hundred-years-old Gothic building]...
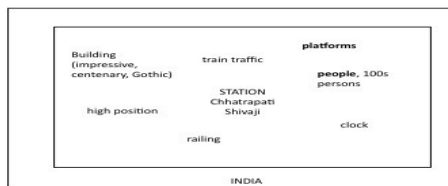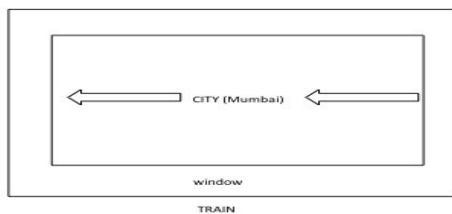

Figure 2: Setting the scene in AD-ES

Jamal looks out of a train window as it races through the city.

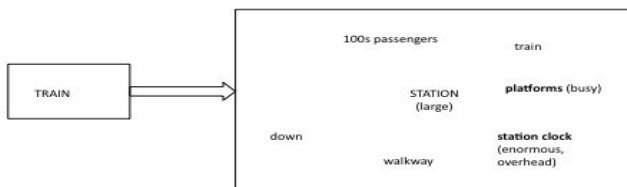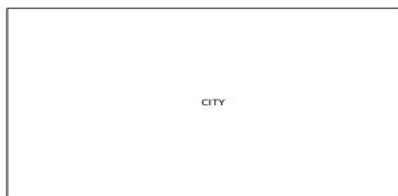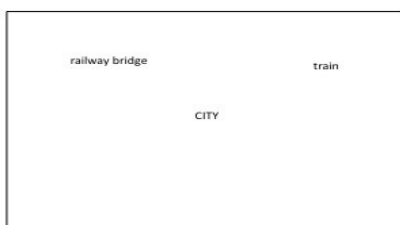The train pulls into a large station...

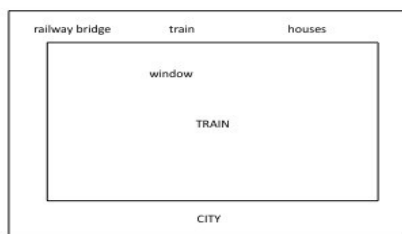Figure 3: Setting the scene in AD-EN

In der Stadt.
[in the city]

CITY

Ein Zug rast über eine Eisenbahnbrücke.
[a train speeds across a railway bridge]

railway bridge   train

CITY

Jamal lehnt sich aus dem Fenster...
[...]

railway bridge   train   houses

window

TRAIN

CITY

Der Zug fährt in einen großen Bahnhof ein...
[...]

railway bridge   CITY   houses
train
window TRAIN   STATION (large)

Hunderte von Menschen...
[...]

window TRAIN
train

wagon door   platform   100s people, masses
downwards   STATION (large)   building (magnificent, colonial style, high)
railing
trains   walkway   clock
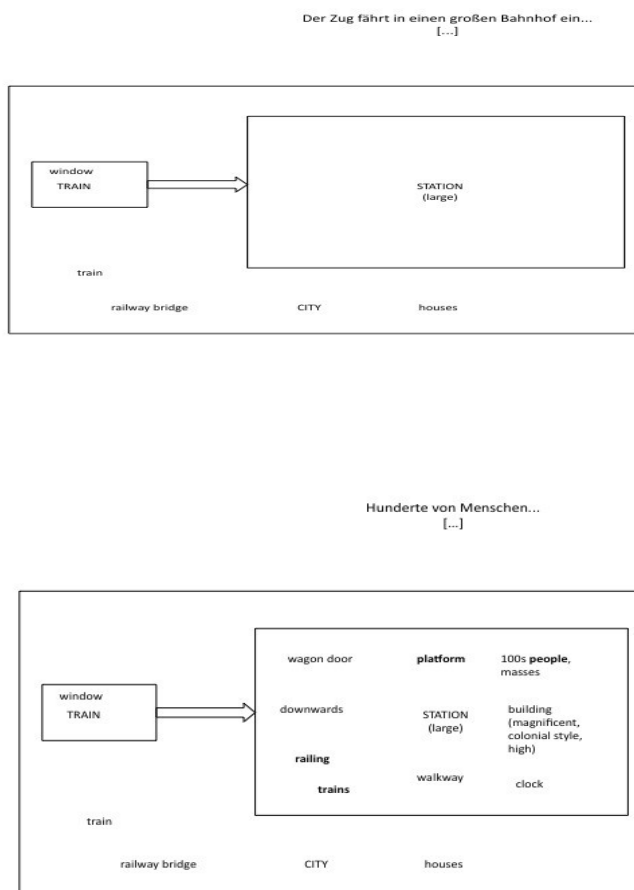railway bridge   CITY   houses

Figure 4: Setting the scene in AD-DE

5. Findings

Having carried out the analysis of the 76 scenes, we found the following types of verbal cueing of setting:

- First, there are clear differences in length. Some audio descriptions leave more time for soundtrack than others. The AD-ES characterises for not leaving much time for the soundtrack, while the other two versions, in

particular the English one, appear to note the relevance of sounds, especially in a film like the one this study analyses, where soundtrack plays a significant role. Auditory cues can also be heard in the background of AD, as our example scene demonstrates (sounds of trains and people in the station). In relation to other narrative entities than setting, auditory cues may also have relevance. In the example scene, Latika's theme is played when Jamal sees Latika in the station, and this is non-diegetic music[3] which does not refer to the setting but portrays a character (Igareda 2012).

- Second, by picking out the references of setting, we noted a certain difficulty in determining whether certain items are part of the setting or whether they are a setting themselves (for instance, a street and a car on that street). This dynamicity emerges in the audio description, for instance when the linguistic representation changes between assigning an item as Figure and as Ground (in the analysed example, TRAIN is variously the frame of reference to Jamal's action as well as an acting subject itself) (see Talmy 2000, Hirvonen 2013).

- Third, regarding strategies of spatial composition, it can be said that both AD-ES and AD-EN use a "component parts" strategy (see Bordwell & Thompson 1990): from details to a big picture (Jamal/train window → train → train station/city). AD-DE begins with the setting and then moves to describe elements in it, i.e. using an 'analytical breakdown' strategy (city → train/railway bridge → Jamal/train window). If we focus on the 'Blickführung' (Schubert 2009) it can be said that AD can point to different parts of setting and thus cue differences in perspective (e.g. train – train window – train station – in front of train station, etc.) (see also Hirvonen 2013). The schematic images provided at the end of the analysis illustrate these different approaches. Comparing the English and the German versions, for instance, we observe that AD-EN provides complex information already with the first utterance, whereas AD-DE composes the setting little by little.

- There are different approaches regarding the generalisation or specification describing some actions, for example: whereas AD-ES refers to the action in the beginning of the scene as *Jamal llega* [Jamal arrives], AD-EN and AD-DE specify this by describing how Jamal looks through train windows at the

---

[3] It is considered an instance of sound "whose supposed source is not only absent from the image but is also external to the story world [the diegesis]" (Chion 1994: 73).

city that passes by. On the other hand, AD-ES can be seen to specify the train station by providing information in a "tourist-guide style" which seems to go beyond the visual appearance (*declared World Heritage site by UNESCO and today it is the station with most traffic in India*).

- The instance in which all three versions describe exactly the same (the clock) demonstrates how the same source text substance is clearly so central that each AD version thematises it.

6. Conclusions

The aim of this study was to describe strategies for the verbalization of setting as well as discuss similarities and differences in different-language audio descriptions. Our first question was related to how the film's setting is verbalized in AD. The analysis carried out shows that 'setting', when understood as place of action (see ʿHandlungsraumʾ in Seiffert 2005), is a complex, multidimensional notion. Our initial definition of setting was that it is the physical, inanimate environment or place in which the narrated action happens. Bordwell and Thompson (1990) term this ʿlocaleʾ (compare also with the notion of ʿUmgebungʾ in Schmitt & Deppermann 2010). This includes not only the broad, container-like aspect of the setting but also all the constitutive objects and other entities in it which belong to the schematic network of the place in question. Thus, a house would represent the broad sense of a setting, where doors, windows, and other things such as garden (depending on the type of house of course) would constitute some typical parts of a house. But tracking cues for setting in the data according to this definition raises questions: How do we decide which elements are ʿconstitutiveʾ and which are not?

Our definition of setting now is: a physical, inanimate place of action having constitutive objects and entities (or things that become constitutive or identifiable with that setting) as well as different levels. These aspects may appear in different ways to the spectator: the setting may be seen in its broad dimension, or our gaze may be directed to certain parts within it (e.g. by changing the viewpoint from a long shot to a close-up). Interesting from the AD perspective is how it verbalises the different perspectives.

Another interesting aspect in respect to defining setting is that, as a place of action, it seems to have multiple dimensions in which action happens. Some of these layers may first appear as objects in the setting (e.g. a car in a street) and then become a place of action themselves, leaving the broader dimension of the setting

in the background. In the film, the setting thus varies from the idea of Mumbai as a city to some very minute parts of it, such as a table in the interrogation room (scene 1, AD-EN) or the screen of a computer (scene 56, AD-DE).

Regarding the second question of this study, which followed how the setting is verbalized in different AD versions (languages) of the same source text, and if these languages/versions reflect shared conceptualizations, we note differences in the global strategy of the three versions, illustrated by our example of analysis. The Spanish audio description uses an explanatory strategy for instance by rendering additional information on the places in which the film is set. The German and English audio descriptions are less explicit in this sense but differ in the style of conceptualising locales and objects (see also Hirvonen 2013): An example can be found in the very beginning of the film when AD-DE describes the place as *a small cranked room* (pointing to the physical appearance) and AD-EN, in contrast, as *interrogation room* (pointing to a social function).

7. Discussion

The analysis reported in this article introduced and demonstrated how three different audio descriptions of the same film construct setting. It was proven that these three different versions provide a perception of the many possibilities and also limits of audio description when verbalising the visual experiences.

After having summarized the main findings and put together the conclusions derived from this study, we also want to cover aspects that were not studied in this study along with everything that can still be part of further research.

Regarding the first aspect mentioned above, the corpus of this study could be analysed so that the construction of the film's different settings (e.g. police station, studio, slum, etc.) are described from the beginning to the end (e.g. what elements are mentioned explicitly, what is implied). It would also be relevant to study what kinds of 'imaginary space' the audio descriptions produce of India.

In the future, because of the distinct conceptualization strategies, it would be interesting to compare AD to the visual source text and to study what aspects lead to choose either the socially or physically profiled strategy. Another interesting topic is that we cannot know whether choosing a certain concept reflects a calculated strategy or rather a non-deliberate naming process (e.g. when culturally unfamiliar objects are seen, are they likely to be conceptualised with a less specific category?).

All in all, our corpus furnishes valuable data for future contrastive and cross-linguistic inquiries.

Maija Hirvonen
University of Helsinki
maija.hirvonen@gmail.com

Paula Igareda
Universitat Autònoma de Barcelona
paula.igareda@gmail.com

Filmography

Slumdog Millionaire (Danny Boyle & Loveleen Tandan, 2008).

References

Bordwell, David (1985). *Narration in the Fiction Film*. London: Methuen.

Bordwell, David & Thompson, Kristin (1990). *Film Art. An Introduction*. New York: McGraw-Hill Companies.

Bourne, Julian & Jiménez, Catalina (2007). "From the visual to the verbal in two languages: a contrastive analysis of the audio description of The Hours in English and Spanish". Díaz Cintas, Jorge, Orero, Pilar & Remael, Aline (eds). *Media for All: Subtitling for the Deaf, Audio Description and Sign Language*. Amsterdam: Rodopi, 175–187.

Braun, Sabine (2007). "Audio Description from a discourse perspective: a socially relevant framework for research and training". *Linguistica Antverpiensia* NS, 6, 357–369.

— (2008). "Audiodescription research. State of the art and beyond". *Translation Studies in the New Millennium* 6, 14–30.

Carroll, Mary & von Stutterheim, Christiane (2007). "Durch die Grammatik fokussiert [Focused through the grammar]". *Zeitschrift für Literaturwissenschaft und Linguistik* 145, 35–60.

Chesterman, Andrew (1998). *Contrastive Functional Analysis*. Amsterdam: Benjamins.

Chion, Michael (1994). *Audio-Vision: Sound on Screen*. New York: Columbia University Press.

Cruse, Alan (2011). *Meaning in Language. An Introduction to Semantics and Pragmatics.* 3rd ed. Oxford: Oxford University Press.

Fryer, Louise (2010). "Audio description as audio drama – a practitioner's point of view". *Perspectives: Studies in Translatology* 18 (3), 205–213.

Grodal, Torben (1997). *Moving Pictures. A New Theory of Film Genres, Feelings, and Cognition*. New York: Oxford University Press.

Hirvonen, Maija & Tiittula, Liisa (2012). "Verfahren der Hörbarmachung von Raum. Analyse einer Hörfilmsequenz [To make space audible: Analysis of a sequence of an audio-described film]". Hausendorf, Heiko, Mondada, Lorenza & Schmitt, Reinhold (eds). *Raum als interaktive Ressource* [Space as interactive resource]. Tübingen: Gunter Narr Verlag, 381–427.

Hirvonen, Maija (2013). "Sampling Similarity in Image and Language – Figure and Ground in the Analysis of Filmic Audio Description". *SKY Journal of Linguistics* 26, 87–115.

Igareda, Paula (2012). "Lyrics *against* images: Music and audio description". *MonTI* 4, 233–254.

Kluckhohn, Kim (2005). "Informationsstrukturierung als Kompensationsstrategie [Information structure as compensation strategy]". Fix, Ulla (ed.). *Hörfilm. Bildkompensation durch Sprache* [Audio description. Image compensation through language]. Berlin: Erich Schmidt, 49–66.

Matamala, Anna & Rami, Naila (2009). "Análisis comparativo de la audiodescripción española y alemana de *Good-bye Lenin* [Comparative analysis of the Spanish and German audio descriptions of *Good-bye Lenin*]". *Hermeneus: Revista de la Facultad de Traducción e Interpretación de Soria* 11, 249–266.

Monaco, James (2009). *How to read a film*. New York: Oxford University Press.

Orero, Pilar (2005). "Audio Description: Professional Recognition, Practice and Standards in Spain". *Translation Watch Quarterly* 46 (1), 7–18.

― (2008). "Three different receptions of the same film. "The Pear Stories Project" applied to audio description". *European Journal of English Studies* 12 (2), 179–193.

Remael, Aline (2012). "For the use of sound. Film sound analysis for audio description: Some key issues". *MonTI* 4, 255–276.

Schmidt-Banse, Hans-Christian (2002). "Die kommunikativen und ästhetischen Funktionen der Filmmusik [The communicative and aesthetic functions of film music]". Leonhard, Joachim-Felix (ed.). *Medienwissenschaft: ein Handbuch zur Entwicklung der Medien und Kommunikationsformen*, *Teilband 3* [Media Studies: A guide for the development of media and communication forms, Volume 3]. Berlin: de Gruyter, 1137–1142.

Schmitt, Reinhold & Deppermann, Arnulf (2010). "Die multimodale Konstitution eines imaginären Raums als interaktive Problemlösung [The multimodal constitution of an imaginary space as an interactive problem solution]". Deppermann, Arnulf & Linke, Angelika (eds). *Sprache Intermedial. Stimme und Schrift, Bild und Ton* [Inter-media language. Voice and text, image and sound]. Berlin: de Gruyter, 199–242.

Schubert, Christoph (2009). *Raumkonstitution durch Sprache*. *Blickführung, Bildschemata und Kohäsion in Deskriptionssequenzen englischer Texte* [Constitution of space through language. Gaze directing, image schemata and cohesion in description sequences in English texts]. Tübingen: Niemeyer.

Schwarz, Monika (1992). *Kognitive Semantiktheorie und neuropsychologische Realität* [Cognitive theory of semantics and neuropsychological reality]. Tübingen: Niemeyer.

Seiffert, Anja (2005). "Räumliches hören. Ein schemaorentierte Analyse der audiodeskriptiven Darstellung der Handlungsräume [To hear space. A schema-oriented analysis of the audio described representation of the space for action]". Fix, Ulla (ed.). *Hörfilm. Bildkompensation durch Sprache* [Audio description. Image compensation through language]. Berlin: Erich Schmidt, 67–86.

Stutterheim, Christiane von & Kohlmann, Ute (2001). "Beschreiben im Gespräch [Describing in conversation]". Antos, Gerd, Brinker, Klaus Heinemann, Wolfgang & Sager, Sven F. (eds). *Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung (HSK)* [Text and speak linguistics. An international handbook of contemporary research]. Berlin: de Gruyter, 1279–1292.

Stutterheim, Christiane von & Klein, Wolfgang (2002). "Quaestio and L-perspectivation". Graumann, Carl F. & Kellmeyer, Werner (eds). *Perspective and Perspectivation in Discourse*. Amsterdam: Benjamins, 59–88.

Talmy, Leonard (2000). *Toward a Cognitive Semantics, Volume 1: Concept Structuring Systems*. Massachusetts: MIT Press.

Vercauteren, Gert (2007). "Towards a European guideline for audio description". Díaz Cintas, Jorge, Orero, Pilar & Remael, Aline (eds). *Media for All: Subtitling for the Deaf, Audio Description and Sign Language*. Amsterdam: Rodopi, 139–149.

Maija Hirvonen holds a MA (University of Tampere) and a PhD (University of Helsinki), both in Translation Studies. Her research activities focus on the latest developments in the audiovisual translation and the newer/less established forms of translating and interpreting. She currently works as a Postdoctoral Researcher at University of Helsinki, Department of Modern Languages and carries out the research project "Multimodal Translation with the Blind (MUTABLE)".

Paula Igareda holds a BA in German Philology (University of Salamanca) and a PhD in Translation (Universitat Pompeu Fabra). She has been a professional translator for many years and has participated in several European projects. Her lines of research are Audiovisual Translation, Accessibility to the Media, and Audio description. She currently works as an associate professor at UAB, UPF and UVic.