punctum.
INTERNATIONAL JOURNAL OF SEMIOTICS

ISSN 2459-2943
https://punctum.gr

06 | 01
2020

Translation and Translatability in Intersemiotic Space

EDITED BY
Evangelos Kourdis & Susan Petrilli

PUBLISHED BY
THE **HELLENIC**
SEMIOTIC SOCIETY

Audio describing the mental dimension of narrative characters. Insights from a Flemish case study.

punctum.gr

ABSTRACT

Audio description (AD) is a service for people with sight loss that makes audiovisual content such as films and TV series accessible to them by verbally describing the visual elements they cannot access. This form of intermodal translation entails various challenges. One of them is how to render orally the emotions, feelings, and other mental states of narrative characters, i.e., elements that we infer from concrete actions, facial expressions, and gestures shown on screen. In practice, we can use various strategies, situated on a continuum ranging from an objective 'describe what you see' approach to more interpretative, subjective descriptions, explicitly naming the mental state underlying the visuals. Although early AD guidelines recommend objective descriptions, recent research has indicated that more subjective approaches may offer various advantages to target audiences in terms of immersion in the story world or imposed cognitive load. In this paper, we present the results of a case study involving the analysis of three episodes from different Dutch-spoken TV series to explore a) what strategies audio describers use to express mental states and b) where do they stand on the objective-subjective continuum. The results show that, contrary to what the guidelines recommend, the descriptions are situated nearer the subjective side of the continuum, suggesting that, when translating visual elements into a verbal form, audio describers tend to look beyond the screen to infer the implicit underlying meaning.[1]

ARTICLE INFO:

Volume: **06**

Issue: **01:2020**

ISSN: **2459-2943**

DOI: **10.18680/hss.2020.0005**

Pages: **85-107**

By: **Bonnie Geerinck** and **Gert Vercauteren**

KEYWORDS:

audio description

intermodal translation

visual-verbal transfer

audio description strategies

character description

[1] The research of this article is part of the CoReAD Project 'Cognitive Research in AD: Towards a model determining the cognitive load of audio described audiovisual products', funded with support from the BOF Research Fund of the University of Antwerp (41/ FA030600/FFB190134).

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

1. Introduction

Audio description (AD) is a service that makes audiovisual products such as films or TV series accessible to people with sight loss. It does so by translating the visual elements in the original work that the target audience does not have access to into a verbal narration inserted between dialogues and other relevant aural components such as music and sound effects (Remael et al. 2015). As such, audio description constitutes an instance of what Kress (2003: 47) calls *transduction*, i.e., an "operation which involves shifts across modes." In translation studies, AD is defined, following Jakobson's (1959) division of translation types, as a form of transmutation or intersemiotic translation (Jakobson 1959: 233). However, with new technological developments leading to new types of translation, we may wonder whether this term is still capable of accommodating all the new multimodal realities that have entered the realm of translation since the introduction of screen communication and audiovisual translation. Kaindl (2013), for example, claims that intersemiotic translation, defined by Jakobson (1959: 233) as "an interpretation of verbal signs by signs of nonverbal sign systems," is somewhat "unfortunate" (Kaindl 2013: 261), given the expansion and diversification of the field.

He suggests moving away from Jakobson's (1959) linguistically based concepts and distinguishing different types of translation based on the concepts of *mode* and *medium*, as perceived by Kress and Van Leeuwen (2001). In this way, we can categorize translations along two dimensions (Kaindl 2013: 261-262). The first dimension looks at the *mode* of the source text and the target text. The mode can be the same, e.g. when we translate a written text in one language into a written text in another language, in which case we have *an intramodal* translation. Or it can be different, e.g., when we translate a written text into a picture book, in which case we have an *intermodal* translation. The second dimension looks at the medium in which the source text and target text are presented. Again, the medium can be the same, e.g. an English novel on paper is translated as a novel on paper in any other language, in which case we have *intramedial* translation. Or it can be different, e.g., when a novel on paper is translated as a play, in which case we have *an intermedial* translation. Given this two-dimensional framework, audio description can be defined as *intermodal* (from the visual mode to the verbal mode) and *intramedial* (e.g., within the medium of film).[2] One of the advantages this classification offers is that it allows for a much more fine-grained analysis of the difficulties arising during the translation process, attributable to either the modal or the medial aspect of the transfer.

[2] We use film here as an example since there are many other audiovisual products to which audio descriptions can be added, such as plays, opera, sports events, museum exhibits, to name but a few. In all these instances, however, the medium of the translation stays the same, and it is only the mode that changes. In other words, all these types of AD are forms of intramedial translation.

In our paper, we take a closer look at the challenges posed by the AD process's intermodal nature, with a particular focus on the description of the psychological (mental and behavioral) dimension of narrative characters. After some theoretical observations and an overview of the existing literature on this particular topic, we will look at how the difficulties inherent to this specific component of the description are handled in Dutch-spoken audio descriptions, based on a case study of three different TV series.

2. Challenges posed by the intermodal nature of audio description

Like any other type of translation, AD is a process that consists of various phases, the main ones being the analysis of the source text and the creation of the target text using appropriate translation strategies. In other words, as Remael et al. (2015) put it, the describer first has to *determine* what elements are eligible for description and then has to *decide* what elements (s)he will eventually include in the description and how. Both these steps in the process are impacted by the specific modes of the source and target text and the AD process's intermodal nature.

2.1. Source text analysis or determining what we can describe

As explained above, audio description involves translating visual images into a verbal narrative for people who do not have access to the visual component of audiovisual products. Therefore, when determining what elements can be included in the description, the describer has to carefully analyze the source text's visual mode. This may seem pretty straightforward, particularly since – contrary to natural languages – we never really have to learn how to 'read' a film to understand it. Our common presumption that we understand the full meaning of visual communication is due to the fact that images, to a higher degree than words, seem to be pre-filled with meaning. Indeed, as Kress (2003: 3) observes: "words are, relatively speaking, empty of meaning, or perhaps better, […] the word is there to be filled with meaning." Whether or not words are indeed empty of meaning remains open to debate. However, it is clear that, compared to images, their meaning is to a significantly higher degree based on convention rather than on resemblance, which has considerable advantages in the case of translation, as will be briefly discussed below. Monaco (2009: 179) makes one important additional observation: "language systems may be much better equipped to deal with the nonconcrete world of ideas and abstractions […] but they are not nearly so capable of conveying precise information about physical realities." In visual communication, the opposite seems to be true. In terms of concrete, physical phenomena, an image "can give us a close nonconcrete reality, it can communicate a precise knowledge that written or spoken lan-

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

guage seldom can" (Monaco 2009: 179). On the other hand, non-concrete, abstract realities such as feelings, thoughts, or other mental states of people, in this case, film characters, are much harder to depict with a high degree of precision and often include a significant implicit component when rendered visually. This point's relevance for source text analysis in audio description is clear, and both the physical and the abstract parts of the visual message are problematic in their way because of this reality.

When analyzing concrete visual elements such as spatiotemporal settings or the physical appearance of characters, a first feature that complicates our choice of what needs describing is the so-called overdetermination or overconcreteness (Schmid 2014) of these elements. Depicted as they look in reality, we cannot pre-select what is relevant and what is not, making it "difficult to discern, whether an element or a property is intentionally represented and therefore belongs to the story or came into the field of vision accidentally" (Schmid 2014: 16-17). This makes it harder to determine whether we need to describe a specific element in an image. A second feature that complicates the analysis is that the meaning of these concrete visual elements is based more on resemblance than convention. This makes recognizing and correctly interpreting a particular element a much more individual, top-down endeavor. Put differently, if the translator does not know a specific word, (s)he can look up its 'conventional' meaning in a dictionary. Suppose, however, that an audio describer does not know a tool shown in a documentary or does not recognize a specific landmark or other cultural references in a film. In that case, it will be much more challenging and sometimes even impossible to arrive at its precise meaning, which will either be lost or only approximated at best.

Settings and physical appearance are far from optional in stories. According to narratological research, however, a sufficient understanding of narrative – including films and TV series, and their audio described versions – depends on insights into the characters' mental dimension, i.e., their actions and reactions, and the feelings, emotions, and other mental states underlying and driving these actions (e.g., Emmott 1997; Fresno 2016; Fresno et al. 2016; Palmer and Salway 2015). Since this mental dimension is a non-concrete constituent, it cannot be precisely depicted in visual communication. Indeed, what can be seen – and is part of the source text analysis – are the characters' facial expressions, gestures, and other actions. Still, the emotions, feelings, or other mental states underlying them, generally remain implicit and will have to be inferred. Just like any other meaning-making process, this inferencing is always an individual and hence subjective endeavor. That means that the result of the analysis of a particular facial expression or gesture may differ depending on the individual describer. In their study of two audio descriptions of the film *The English Patient* (Minghella 1996), Palmer and Salway (2015: 134) discuss a scene in which a particular mental state of one of the characters is described in two very different, even opposite, ways, which highlights the difficulty of interpreting and rendering the meaning of non-concrete visual information.

And, as the authors rightfully conclude, 'small differences in descriptions […] can result in big potential differences for an audience's understanding of the story' (Palmer and Salway 2015: 135). As will become apparent in the following section, the audience's experience is not only influenced by differences in what is described. How descriptions are formulated is equally important, not only for the overall understanding of the story but also – as recent research suggests – for the cognitive load imposed on the target audience and hence on their enjoyment of the audio described product.

2.2. **Target text creation or how to formulate the description**

Once we have determined what elements are eligible for description, two decisions will have to be made in creating the target text. First of all, the describer has to decide which of these elements he will include in the description: since audio descriptions have to fit in between dialogues and cannot interfere with other sound effects – the meaning-making channels people with sight loss do have access to – it will often not be possible to include all relevant information and a selection will have to be made (Fryer 2016; Remael et al. 2015; Vercauteren 2012). Again, this can be ascribed to the fundamentally different logics governing the visual and verbal mode: while in visual communication, information is presented simultaneously and can be taken in holistically, in verbal communication, information has to be presented and processed sequentially. Therefore, only a small portion of a film's wealth of visual elements can be rendered in the short timespan between two dialogues. This explains why most early guidelines and research in audio description (e.g., Snyder 2005; Vercauteren 2007; Kruger 2010; Rai et al. 2010) focus on designing protocols and methodologies for content selection and helping describers to determine what is relevant and what not.

Once the describer has selected the elements (s)he will include in the description, (s)he has to decide how to verbalize them. When it comes to the style of audio description or, how to describe, the main bone of contention has always been whether to describe objectively or subjectively. Even though one may wonder whether 'objectivity' is at all possible in a highly individualized process such as meaning-making, many AD guidelines state that describers should only describe what they see without any interpretation (e.g., Benecke and Dosch 2004; Snyder 2010), or that "[t]he description must be given objectively in order not to impose the describer's own feelings but rather provoke the listener's" (Rai et al. 2010: 61). This rule of objectivity is particularly stressed when it comes to the audio description of facial expressions or gestures, probably not surprising given the inherently implicit dimension present in this kind of non-concrete information when presented visually. By describing the physical traits visible on the screen, the describer avoids rendering emotions or feelings that the author did not intend (cf. the example from *The English Patient*, mentioned above). Two observations have to be made in this respect.

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

First of all, this rigid opposition between objective and subjective descriptions does not reflect the far more complex and nuanced reality of communication, in general, and audio description in particular. Most of the descriptions of characters' facial expressions and gestures or the emotions, feelings, or other mental states cannot be categorized as purely objective or subjective, which has been acknowledged in recent AD research. Adopting a narratological approach to audio description, Kruger (2010) proposes to move away from precisely describing what we see on the screen and instead render the narrative effect of this visual information. Essential in this account of what he calls 'audio narration' (Kruger 2010: 233) is that the concrete verbalizations will be located on a continuous scale ranging from objective description to a narration that enhances the audience's immersion in the story world. A similar claim is made by Palmer and Salway (2015: 131), who state that in audio description, "[t]here is a continuum rather than a simple dichotomy." The continuum they refer to has a philosophical basis and is called the thought-action continuum (Palmer 2004). Descriptions such as "she shrugs her shoulders" would be situated nearer to the action end of the continuum, and descriptions such as "she doesn't care" are nearer to the thought end. Descriptions such as "she shrugs indifferently" could then be placed more in the middle. The corpus analysis done by Palmer and Salway (2015) does indeed show that audio description is far from 'either-or' and much more a matter of degree, with most descriptions situated towards the middle of the continuum.

A second, and probably more critical observation, is that the objectivity advocated in early guidelines may often not be the most desirable choice. First of all, objective descriptions tend to be longer than subjective ones (Vercauteren and Orero 2013). Interestingly enough, we find this point also in the German guidelines (Benecke and Dosch 2004), which acknowledge that given the shortness of the pauses between dialogues, short (interpreted) descriptions such as "he looks tensely" should be preferred over longer (objective) ones such as "he has squeezed his eyes shut." Second, research has shown that descriptions situated nearer the subjective or interpreted narration or thought end of the AD continuum are beneficial for the target audience. Walczak and Fryer (2017) carried out an experiment in which they tested two different descriptions, an objective and a creative one, on an audience of people with sight loss and found that the creative description improved the participants' immersion or presence in the story world, creating a more intense narrative experience. In a similar vein, Fresno et al. (2016) tested participants' memory with two different types of descriptions of the physical appearance of characters, which they termed 'visual' (i.e., objective) and 'semantic' (i.e., subjective). They found that the latter yielded better results and hypothesized that the participants tried 'to go through the merely visual descriptions and extract their semantic meaning' (Fresno et al. 2016: 160). They suggest that, as a result, these story world semantic ADs may result in a lower cognitive load imposed on the target audience. This observation forms the starting point of the Belgian research project CoReAD:

'Cognitive Research in AD – Towards a model determining cognitive load in audio described audiovisual products.' The main goal of this project is to investigate what elements in an audio description influence cognitive load. More specifically, it explores how different formulations of narrative characters' mental states impact this load. In the first phase of the project, which is a smaller-scale replication of Palmer and Salway's (2015) research, existing Dutch-spoken audio descriptions were analyzed to see how the characters' mental dimension is described and to determine whether we can identify differences in AD approaches (objective vs. subjective) to specific intra- and extratextual parameters of the audiovisual product described. The analysis and its result, to be discussed in the remainder of this paper, will form the basis of an experiment aiming to test whether ADs situated on different positions on the continuum do indeed result in different cognitive load, as suggested by Fresno et al. (2016).

3. Methodology

For our case study, we analyzed three audio described Dutch-spoken TV series broadcast between 2013 and 2019: *Keizersvrouwen* (Sombogaart and Vos 2019-2020), *Beau Séjour* (Basteyns and Beels 2017), and *Eigen Kweek* (Vanhoebrouck 2013-2019). For each series, we randomly selected one episode, each with a running time of 50 minutes. Relevant to the analysis is that the word count of the three audio descriptions is different: 3187 words for *Keizersvrouwen*, 2330 words for *Eigen Kweek*, and 4248 words for *Beau Séjour*. We chose these TV series in particular as we believe they are a good representation of the Belgian-Dutch mediascape and because they are one of the few Dutch-spoken TV series that have an audio described version available. They were also selected to represent three different genres: *Keizersvrouwen* is a crime series, *Beau Séjour* is a mystery-drama series, and *Eigen Kweek* is a comedy series. Various studies and guidelines (e.g., Fryer 2016; Morisset and Gonant 2008; Ofcom 2017; Rai et al. 2010) have expressed the need for AD and AD's style to change according to film genre. We decided to work with different genres for this exact reason, namely, so that we could consider whether there are differences in description strategies that can be ascribed to the specific genre of the series.

To better understand how narrative characters' mental dimension is commonly described, we looked at the specific language used in audio descriptions to present this kind of information. To do so, we determined the most frequently used formulations and then classified them according to their syntactic and semantic properties. We chose this approach because it allows for a systematic analysis of the degree of interpretation according to the thought-action continuum proposed by Palmer and Salway (2015). We first transcribed the actual audio descriptions, and then analyzed and categorized them according to their linguistic properties.

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

The first step in this process was to record all the verbs indicating any action or reaction performed by the characters. Secondly, all the words and phrases signifying emotion or thought were marked. After this preparatory work, we conducted a linguistic analysis of the phrases that express behavioral and mental states. The final step was to analyze the descriptions in terms of their position on the thought-action continuum to infer their degree of subjectivity. In the section that follows, we will first discuss the different formulation strategies for descriptions of characters' actions, which will also be referred to as *action descriptions*.

3.1. Classification of action descriptions

In general, the behavioral dimension of characters in a film or TV series is understood as their actions and reactions, namely everything characters do through their physical movements, which are visible on screen (Palmer and Salway 2015; Vercauteren 2014). Existing AD research (Henkens 2014) shows that characters' actions are usually described according to four different structures, which are: a) the use of a standard verb with no complement of manner; b) the use of a troponym as an alternative for these standard verbs; c) the use of a verb in combination with an adjunct of manner; and d) the use of figurative language, such as similes or metaphors. The table below illustrates these four different strategies employing examples from our corpus, all referring to the action of walking.

Table 1. **Examples of formulation strategies for describing characters' behavioral dimension**

A. **standard verb**	B. **troponym**	C. **verb + adjunct of manner**	D. **figure of speech**
Een vrouw in het wit **loopt** door een imposante inkomhal. A woman dressed in white **walks** through an impressive entrance hall. *(Keizersvrouwen)*	**Luc ijsbeert.** Luc **paces up and down**. *(Beau Séjour)*	Xandra **loopt blootvoets** over een smal muurtje. Xandra **walks barefoot** on a narrow wall. *(Keizersvrouwen)*	Ze wendt haar bebloede gezicht naar de deuropening, werpt nog een blik op haar lijk in de badkuip en **maakt zich uit de voeten**. She turns her bloody face to the doorway, takes another look at her body in the bathtub, and **takes to her heels**. *(Beau Séjour)*

The first strategy is the most basic way to describe narrative characters' actions: it is straightforward, but it usually doesn't say much about the specific manner in which the action is carried out. Here, we use the term 'standard verb' to refer to simple action verbs, without any ambiguous meaning attached to them, such as the verbs: *to look, to walk, to sit*, etc. They only present the act itself, and do not consider the characters' underlying mental state or motivation. The second strategy is similar to the first, as concerns its syntax structure, but different in its semantic nature. Troponyms are "verbs that express a particular manner of doing something" (Salway 2007: 22), which is why they are especially useful in AD: they provide a solution to the dilemma between the time constraints discussed above and the need for specificity in rendering visual information. Besides, they offer welcome alternatives for standard verbs and help avoid repetition and monotony in the descriptions. The third strategy consists of combining a verb with an adjunct of manner, such as an adverb, an adverbial adjective, or a relative clause. Comparable to the second strategy, this type of structure is used by audio describers to enhance the meaning of the action, which makes the AD more engaging for the audience. In our classification system, *adjuncts of manner* are defined as words or phrases that provide additional information and insight regarding how an action is carried out. Verbs combined with adverbs of place or time, for example, are not included in this category. The final strategy is somewhat different from the ones mentioned since it is not defined by a specific syntax structure but by how meaning is conveyed, namely through figurative language. Figures of speech, such as similes and metaphors, are useful stylistic devices, as they enrich the vocabulary of audio descriptions.

In addition to characters' actions, audio descriptions also describe characters' motivations, feelings, and thoughts, which we will call *thought descriptions.* The following section presents an overview of the different ways these thought descriptions can be presented by audio describers and introduces a classification model consisting of five different formulation strategies.

3.1. Classification of thought descriptions

According to Salway (2007), one of the main kinds of information provided by audio descriptions is the emotional state of characters. In the present study, this 'emotional state' is understood in a broader sense than only a character's emotional properties. It refers to what Margolin (2007) calls the mental dimension of characters, namely their perceptual, emotive, volitional, and cognitive properties. In other words, it concerns everything a character observes, feels, wants, and thinks. The mental dimension is the character's 'interiority' which entails 'inner states, knowledge and belief sets, traits, intentions, wishes, dispositions, memories, and attitudes' (Margolin 1990: 844).

The computational corpus analysis carried out by Salway (2007) revealed that audio describers usually convey this information by using one of the following formu-

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

lations: a) the use of a verb modified with an adverb; b) the use of the verb *to look* (as in *to seem*) followed by an adjective; and c) the use of action descriptions involving the character's head, face, or eyes. However, when analyzing the corpus, we discovered some other interesting formulation tendencies, of which Salway takes no special notice. First, in his analysis, it appears as if an adverb's addition is always necessary to convey emotion through verbs in AD. However, in our corpus, we also found instances in which the use of an action verb alone is enough to convey the mental state behind that action. An example of this is the audio description 'Ook Charlie moet lachen' [Even Charlie has to laugh] (*Beau Séjour*) in which the verb *to laugh* indicates an emotion and does not need to be modified with an adverb to do that. Secondly, Salway does not acknowledge the use of figurative language as a way to describe emotional states. This is probably because figures of speech do not have a fixed syntax structure, making it difficult to detect them solely through a data-driven computational analysis (Semino 2017; Stefanowitsch 2007). Although manual analysis has undeniable disadvantages, it enables us to detect such figures of speech. This is an advantage in audio description since these figures are used to be creative and vivid, enriching the AD language to better suit the narrative style of the film or TV program to be described. Therefore, we treat figurative language as a separate description strategy in this study.

Table 2. **Examples of formulation strategies for describing characters' mental dimension**

A. **single verb**	B. **verb + adverb/phrase**	C. **description + part of the head/face**	D. **(verb reflecting an impression +) adjective**	E. **figure of speech**
Kristel **huilt** en **roept**. Kristel **cries** and **screams**. *(Beau Séjour)*	**Woedend slaat** Xandra een pot van de tafel. Xandra **furiously knocks** a pot off the table. *(Keizersvrouwen)*	Julita **rolt met de ogen**. Julita **rolls her eyes**. *(Eigen Kweek)*	Xandra **lijkt** even verrast, maar stapt dan doelbewust op de meisjes af. For a moment, Xandra **looks surprised**, but then resolutely approaches the girls. *(Keizersvrouwen)*	Bernard **schrikt zich een hoedje** wanneer Patrick plots achter hem staat. Bernard **jumps out of his skin** when Patrick suddenly stands behind him. *(Eigen Kweek)*

Based on these observations, we would like to suggest a more detailed classification of how information about characters' mental states can be described, namely: a) the use of a single verb implying an emotional state; b) the use of a verb modified with an adverb or phrase; c) descriptions that mention a part of the head or face; d) the use of an adjective or phrase indicating emotion or thought, either with or without a verb that reflects an impression such as *to look*, and e) the use of figures of speech. Table 2 presents an overview of our analysis's five categories, again with examples from our corpus.

The first strategy in this classification offers one of the most concise ways to describe characters' mental states, as it only requires one word to do so. Some of the most commonly occurring verbs in this category are *to smile, cry, startle,* and *frown.* Similar to the first two strategies used for action descriptions, namely the use of common verbs and troponyms, this type of description is a convenient way to manage time constraints. The second strategy requires a somewhat longer description since the verb is modified with an adverb or phrase. In this case, the verb usually refers solely to the action, and the adverb or phrase refers to the mental state behind that particular action. The third strategy for describing characters' minds consists of phrases in which parts of the head or face are mentioned, such as the eyes, mouth, eyebrows, or the head and face in their entirety. The fourth category consists of sentences in which an emotional state or reasoning is implied without any reference to action. This kind of phrase is usually formed with the verb *to look* – or another verb that reflects an impression, such as *to appear, to* be, plus an adjective or phrase. Another typical composition is describing a character solely with an adjective without the use of a verb. The final strategy consists of using figures of speech that express a particular emotional state. Like their counterparts in the categorization of action descriptions, these descriptions offer a stylistic way to make the language of the AD more varied. Although these five categories are not always mutually exclusive – as is often the case in linguistic analyses – we decided only to ascribe one category to each description.

As the classifications of the action and thought descriptions presented above show, visual representations of narrative characters' mental states can be rendered verbally in very different ways. Some of the underlying reasons are intratextual, i.e., sometimes there is not enough time for a detailed description, and a short alternative has to be chosen for reasons of economy. Other reasons are extratextual and may be ascribed to the describer's individual preference for one formulation over another. What describers have to be aware of is that this personal preference also entails differences in the level of their description's subjectivity. One of the advantages of the classification presented above is that it allows us to indicate this subjectivity level based on the thought-action continuum, as suggested by Palmer and Salway (2015).

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

3.2. **Thought-action continuum**

The thought-action continuum consists of four different ways in which we can describe the characters' mental state or, in other words: "the extent to which the describer makes explicit for the AD audience what the sighted audience is expected to infer from the moving image" (Palmer and Salway 2015: 136). In order from action-oriented to thought-oriented, they are the following: a) descriptions of simple actions, such as movements of the body or face without explicitly referring to the underlying psychological state; b) descriptions of facial expressions; c) modified action descriptions which provide insight into how the action is performed through the use of a troponym, an adverb, or a phrase; and d) descriptions that make statements about the mental state characters appear to be in (Palmer and Salway 2015: 136). Table 3 illustrates these categories through some examples from the corpus of the present study.

Table 3. **Examples of descriptions according to their position on the thought-action continuum**

A. **simple action** (action end of the continuum)	B. **facial expression** (midway point: action-oriented)	C. **modified action** (midway point: thought-oriented)	D. **apparent mental state** (thought end of the continuum)
Xandra **huilt**, Michiel loopt weg van een persmoment. Xandra **cries**, Michael walks away from a press briefing. *(Keizersvrouwen)*	Pepita **fronst**. Pepita **frowns**. *(Eigen Kweek)*	Het meisje **staat als aan de grond genageld**. The girl **is riveted to the spot**. *(Beau Séjour)*	Xandra **lijkt niet helemaal op haar gemak**. Xandra **does not seem to be entirely at ease**. *(Keizersvrouwen)*

The categories of our classification and those of Palmer and Salway's (2015) classification are not entirely equivalent, as they serve a different research aim. The former allow us to analyze how audio descriptions of a character's mental dimension are formulated; the latter enable us to investigate to what extent this non-concrete narrative constituent's implicit dimension is rendered explicit, i.e., they offer an insight into the degree of the description's subjectivity. However, they do complement each other. In general, using single verbs expressing emotions and thoughts, i.e., the first strategy in our classification, corresponds to more action-oriented descriptions, namely those describing simple actions and facial expressions. Descriptions such as 'Xandra huilt' [Xandra cries] (*Keizersvrouwen*) and 'Kato glimlacht' [Kato smiles] (*Beau Séjour*) illustrate this. Since descriptions in which the head, face, eyes, eyebrows, and mouth are men-

tioned usually refer to facial expressions, the third strategy in our classification generally falls on the action side of the continuum as well. The second strategy in our classification is usually consistent with what Palmer and Salway (2015) call 'modified action descriptions,' as they entail the combination of a verb with an adverb or phrase and are thus located somewhere in the middle of the continuum. Descriptions created according to the fourth and fifth strategy (descriptions in which a verb expressing an impression, an adjective, or a figure of speech are used) generally correspond to descriptions on the thought end of the continuum.

4. Results and discussion

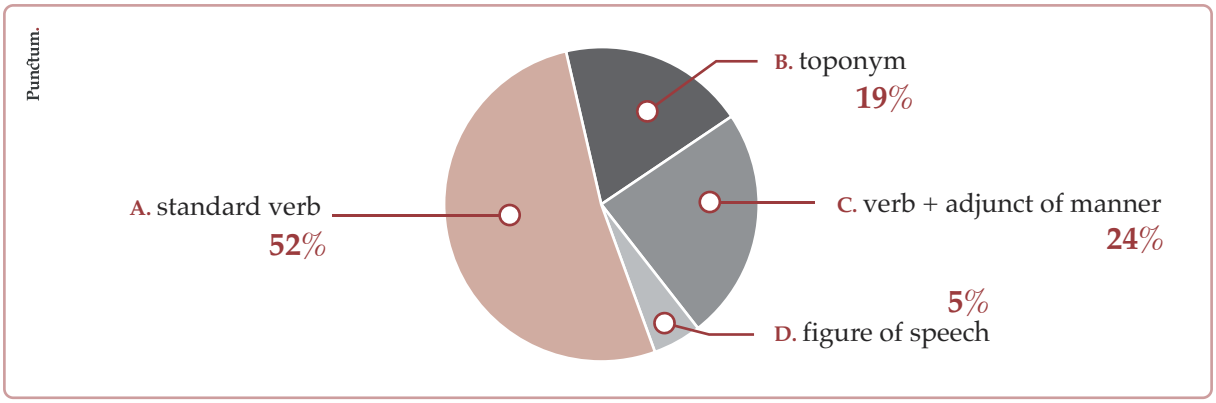
In the previous section, we presented two different classification models for linguistically analyzing audio descriptions: one for the description of characters' actions and the description of characters' emotions and thoughts. In addition to this linguistic analysis, our study also tries to correlate these with varying degrees of subjectivity or explicitation, based on Palmer and Salway's (2015) thought-action continuum. In the previous section, we presented a model to analyze action and thought descriptions regarding their linguistic formulation and degree of subjectivity. This model was tested in a small-scale case study to test its applicability and, at the same time, to get a first idea of how narrative characters' mental states, as expressed through their actions and thoughts, are currently being described in Dutch-spoken AD. We were particularly interested in finding out whether some formulations were used more than others and – in the light of the recently presented advantages of more subjective formulations – where descriptions are located on the thought-action continuum.

4.1. Analyzing the audio description of characters' behavioral dimension

As explained in section 3.1, characters' actions are usually described according to four different strategies: a) the use of common verbs; b) the use of troponyms; c) the use of a verb in combination with an adverb, adverbial adjective, or relative clause which indicate the manner of the action; and d) the use of figures of speech. To determine how the behavioral dimension is described at present, we analyzed all the action descriptions in the corpus and calculated each category's frequency. The results of this analysis are shown in the pie chart below. In addition to a general analysis of the entire corpus, we also conducted a separate study of each episode. We then compared the results to investigate whether the audio descriptions of the three TV series show individual differences in terms of formulation strategies for action descriptions. The comparative analysis did not reveal any significant differences, so we will not discuss the results here.

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

Figure 1. Formulation strategies for describing characters' behavioral dimension, expressed in percentages according to their frequency



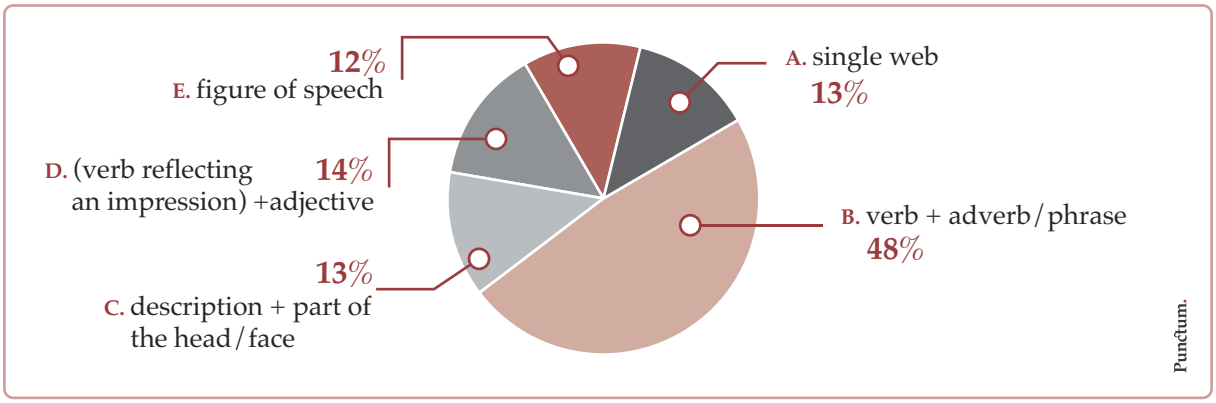
The chart demonstrates that using standard verbs is the most frequently occurring strategy: more than half of all the action descriptions are described in this way (52 percent). This strategy's predominance can be attributed to the time constraints the logic of verbal communication imposes on audio describers, as it forces them to opt for shorter formulations and a less in-depth presentation of the characters' actions. Another explanation could be that actions are a narrative constituent presented in a concrete form in visual communication, i.e., they are directly visible and can more easily be 'named' than non-concrete realities. However, the pie chart shows that the other (albeit smaller) half of the actions in the corpus are described with either the use of a verb modified with an adjunct of manner (24 percent), the use of troponyms (19 percent), or the use of figures of speech (5 percent). These strategies have in common that they enrich the meaning of the action since they provide more details concerning the action's motivation. For instance, troponyms are often indicative of the underlying feelings and intentions of the action described. By using a troponym, audio describers hint at the character's mental state without explicitly mentioning a particular emotion. For example, the verb *'to admire'* in the audio description 'In de hal van de flat bewondert ze enkele zwart-witfoto's van jongedames in lingerie' [In the hallway, she admires some black and white photographs of young women in lingerie] (*Keizersvrouwen*) not only indicates that the woman is looking at something, but at the same time it implies that she appreciates what she is looking at. The same accounts for adjuncts of manner that are added to an action verb to provide more information about either the physical way in which something is done (e.g. 'Xandra loopt blootvoets over een smal muurtje' [Xandra walks barefoot over a narrow wall] (*Keizersvrouwen*)), or the character's emotional state (e.g. 'De inspecteur loopt boos weg' [The inspector walks away angrily] (*Keizersvrouwen*)). An interesting finding in this respect is that descriptions in which the mental state is expressed through an action verb comprise more than one out of four (27 percent) of a total of 1196 action descriptions in the corpus. Overall, these results suggest that concrete visible actions and non-concrete invisible minds in film often in-

tertwine in their verbal description, reflecting Palmer's (2004: 212) argument that 'the mind extends beyond the skin.' However, in addition to expressing emotions through actions, thought descriptions can also be presented using various other strategies.

4.2. Analyzing the audio description of the characters' mental dimension

Concerning thought descriptions, we introduced the following categorization: a) the use of a single verb; b) the use of a verb plus an adverb or phrase; c) descriptions which mention part of the head or face; d) the use of an adjective or phrase, either with or without a verb that reflects an impression; and e) the use of figurative language. To assess how the mental dimension of audio described characters is represented, we classified all the corpus descriptions according to the strategies above and calculated their frequency, both for the corpus in its entirety and the three TV series separately. The former's breakdown is shown in the pie chart below, while the latter is illustrated in Figure 3.

Figure 2. Formulation strategies for describing characters' mental dimension, expressed in percentages according to their frequency



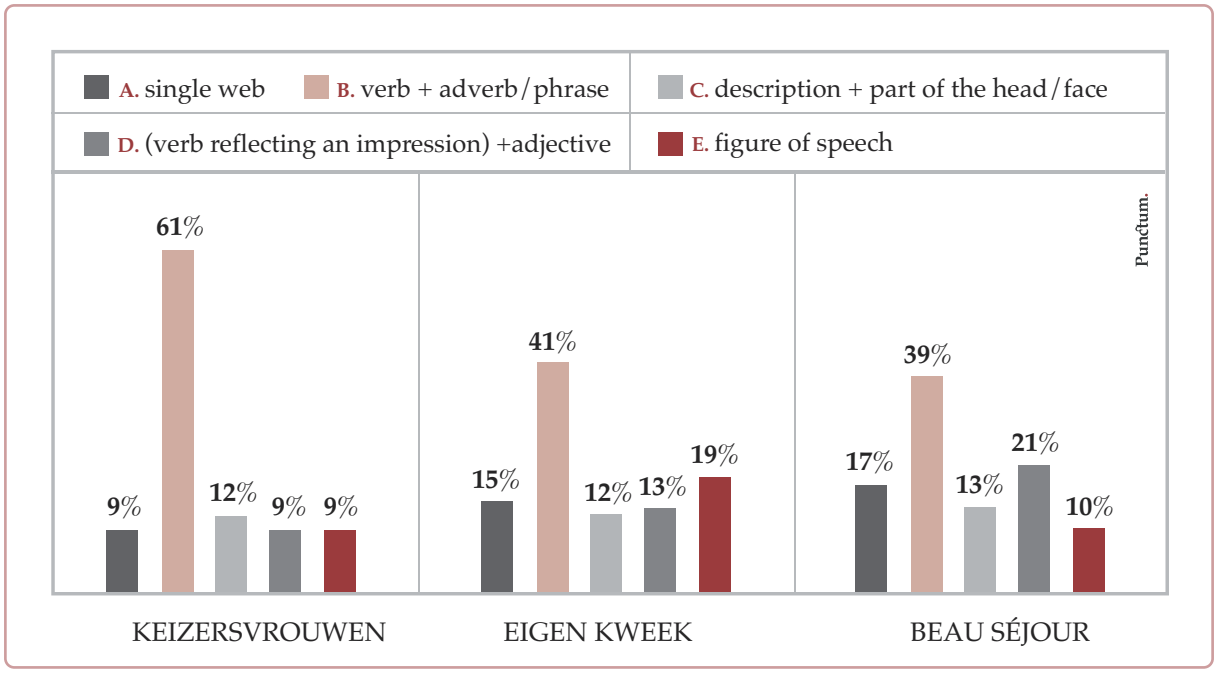
As the data shows, thought descriptions which are composed of a verb combined with an adverb or phrase are the strategy that is used most often to audio describe the mental dimension of characters: 191 out of a total of 400 thought descriptions in the corpus were described according to this strategy (48 percent). This is almost as much as all the other strategies combined, which all occur in 12 to 13 percent of the descriptions. One possible explanation of the predominance of the verb + adverb/phrase strategy could be that the implicit dimension inherent in the visual representation of characters' mental states is not sufficiently explained by a verb alone, i.e., that adverbs name emotions and feelings more explicitly than verbs do and are therefore added in the description. Another explanation could be what we already pointed out in the previous section, namely that the behavioral and mental dimensions are closely linked in

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

audio descriptions. Hence, characters' emotions and thoughts are often expressed through the description of an action. To confirm this hypothesis, we calculated how often thought descriptions across all categories are conveyed through action verbs and found that – whether through the use of a single verb, a verb plus an adverb, or a figure of speech – 81 percent of all 400 thought descriptions in the corpus is described in this manner. This finding seems to corroborate Palmer and Salway's (2015) argument against the supposedly clear-cut dichotomy between a character's mental state and the physical characteristics which reflect that state.

However, a more fine-grained comparison of the three different series shows some marked differences. First, the crime series *Keizersvrouwen* has a significantly higher percentage (61 percent) of thought descriptions of the second category, namely the use of a verb plus an adverb or phrase, than the other two series, which both have around 40 percent. Second, the audio descriptions in the TV series *Eigen Kweek* show a higher percentage of figures of speech than the others: 19 percent for *Eigen Kweek* compared to 9 percent for *Keizersvrouwen* and 10 percent for *Beau Séjour*. Finally, the use of verbs expressing an impression combined with an adjective was found to occur more frequently in the drama series *Beau Séjour* (21 percent), on average 10 percent more than in *Keizersvrouwen* (9 percent) and *Beau Séjour* (13 percent). It seems possible that these results are due to the genre differences between the three series. The audio describer of *Keizersvrouwen* might have opted for more action-oriented descriptions of the characters' mental state because crime series generally have more action scenes.

Figure 3. Formulation strategies for describing characters' mental dimension, expressed in percentages according to their frequency, for the different TV series



In contrast, *Eigen Kweek* is a comedy containing a lot of verbal humor, which may explain the high number of figures of speech in the audio description, since this offers a creative and succinct way to paint a picture of the general atmosphere. Whereas audio describers might be cramped for space to provide in-depth information about characters' emotions and thoughts in TV series filled with action and dialogue, such as in *Eigen Kweek* and *Keizersvrouwen*, drama series usually leave more room for longer descriptions of this kind. Fryer (2016: 119) highlights the same notion, saying that "description for action movies is best kept short and punchy," while descriptive utterances for non-action movies usually "feature longer sentences, with a more lyrical turn of phrase." These results are reflected in the word count of the audio descriptions, which was by far the highest for *Beau Séjour* (4248 words) and the lowest for *Eigen Kweek* (2330 words). However, this is only a hypothesis, and further investigation into the association between genre and AD is required to confirm it.

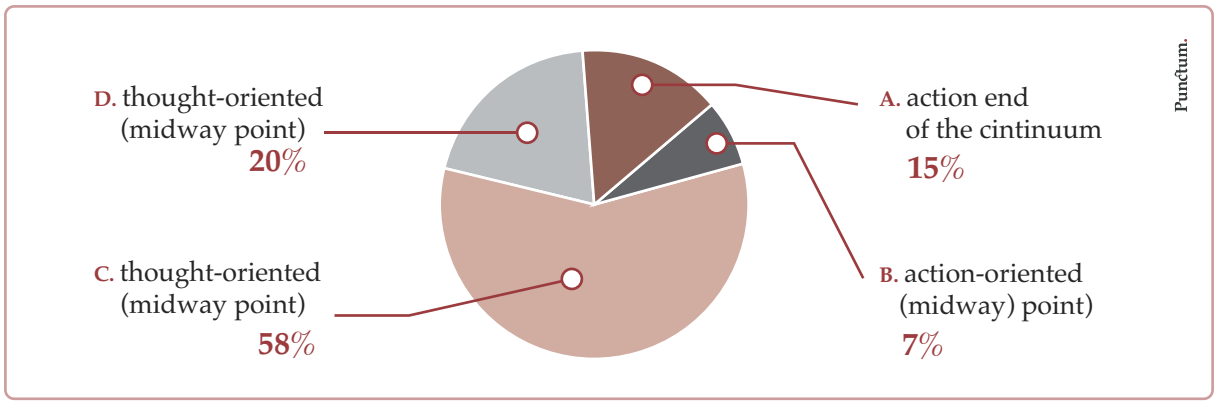
To conclude this section, we would like to point out that out of the total of 1033 sentences in the concise400 sentences (39 percent) include a reference to characters' emotions or thoughts, reflecting the importance of insights into characters' mental states to be able to follow and understand the story. In the final section of our analysis, we will look at where these mental states' descriptions are on the thought-action continuum to gauge their level of subjectivity.

4.3. Degree of Interpretation

As pointed out in the literature review, audio describers are generally advised against explicitly naming the emotions or thoughts of a character, and encouraged to describe the physical characteristics used to express these emotions and thoughts. To assess whether or not this recommendation truly manifests itself in reality, we analyzed how information about the characters' mental dimension is conveyed at present. As a basis for this analysis, we used Palmer and Salway's (2015) thought-action continuum, which places descriptions of characters appearing to be in a certain mental state at the thought end of the spectrum, and descriptions that only mention the physical characteristics of this mental state at the action end of the spectrum. In the middle of the continuum, descriptions which give information about both the action itself and the mental functioning behind the action, such as motivations, intentions, attitudes, and feelings, are positioned. The pie chart below illustrates where the audio descriptions in our corpus are placed on the continuum by giving each approach's percentage frequency.

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

Figure 4. Position of descriptions on the thought-action continuum, expressed in percentages according to their frequency

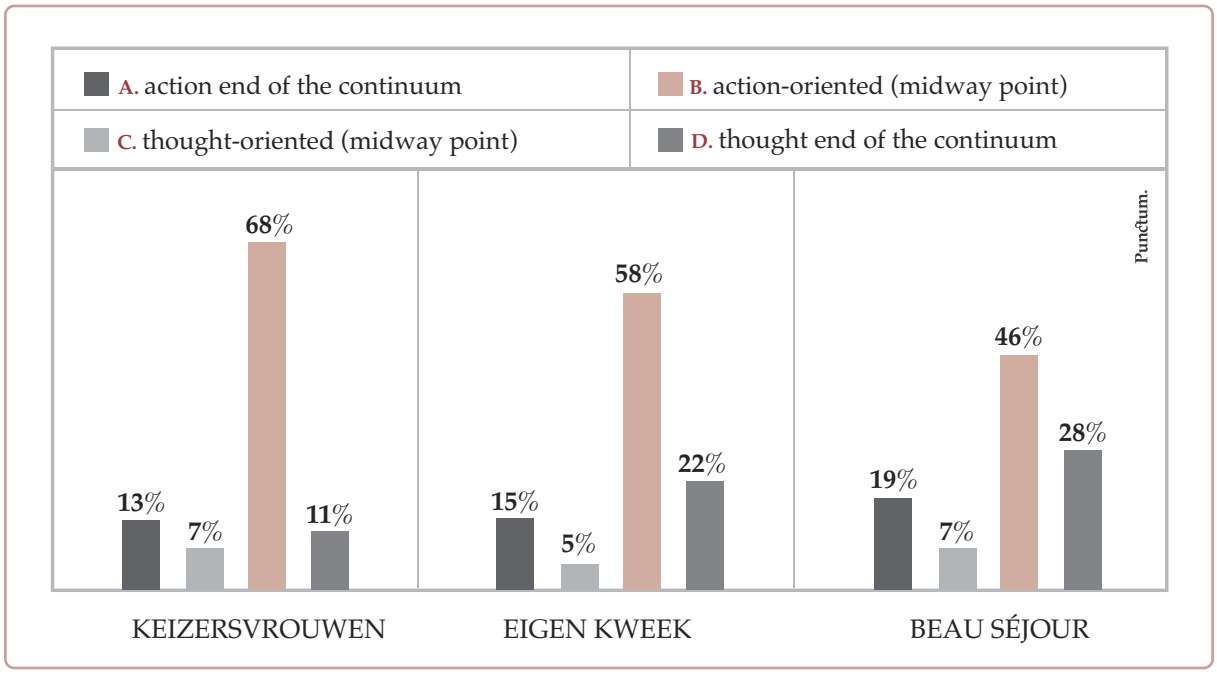


From the chart data, we can infer that information about the characters' mental dimension is primarily conveyed through descriptions that fall midway on the continuum, namely descriptions modified with an adverb, troponym, or phrase, which make up 58 percent in total. This finding is consistent with that of Palmer and Salway (2015), who observed a general tendency towards these kinds of descriptions in audio description for British film and TV. This suggests that describing the mental dimension of characters usually entails some form of interpretation on the part of the audio describer, because (s)he has to render the implicit visual elements into an explicit verbal form, and therefore infers the underlying meaning while at the same time still staying true to what can be seen on screen (i.e., the actions). Interestingly, descriptions with the second-highest percentage frequency were those nearest the continuum's thought end (20 percent). In contrast, descriptions of simple actions (15 percent) and facial expressions (7 percent) were found to occur least frequently. This outcome is somewhat surprising since it goes against what various AD guidelines, such as the Spanish (AENOR 2005), German (Benecke and Dosch 2004), Greek (Georgakopoulou 2010), French (Morisset and Gonant 2008), and American guidelines (Snyder 2010), usually recommend when it comes to describing characters' mental state. Hence, this observation may support the hypothesis that describing as 'objectively' as possible is not always a desirable or possible description strategy. Some degree of interpretation will occur in most audio descriptions of the characters' mental states.

Another interesting finding resulted from the comparison of the audio descriptions of the three different TV series. In the chart below, we can see that descriptions on the thought end of the continuum – thus, more subjective descriptions – is much higher in *Beau Séjour* (28 percent) than in *Keizersvrouwen* (11 percent). Interestingly, the use of descriptions on the midway point of the continuum – that is, thought-oriented descriptions which provide insight into how the action is performed – is highest in *Keizersvrouwen* (68 percent) and lowest in *Beau Séjour* (46 percent). Although these re-

sults need to be interpreted with caution, it may be possible that audio describers of drama series are inclined to a more interpretative, subjective approach, and depend less on action verbs when describing characters' psychological dimension, and that audio describers of action series, on the other hand, lean toward a more moderate approach: both the (objective) action that represents the character's mental state and the (subjective) underlying meaning of that action are included in the description. These observations further support the possible association between film genre and description strategies, and might even indicate that when the genre of a series is more action- or thought-oriented, the accompanying audio description will also be more action- or thought-oriented.

Figure 5. Position of descriptions on the thought-action continuum, expressed in percentages according to their frequency, for the different TV series



5. Conclusions and further research

The present paper started with the observation that audio description, as an instance of intermodal translation, presents various challenges related to the modal differences between the audiovisual source text and the verbal target text and to the intermodal nature of the translation process itself. One constituent that proves to be particularly problematic in this respect is the psychological dimension of narrative characters expressed through their actions, facial expressions, and other body movements. Since this dimension is implicit when expressed visually, rendering it verbally always involves

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

some degree of explicitation and (subjective) interpretation. In our case study, we wanted to gain an insight into the different strategies currently used in Dutch audio descriptions and see whether these are more objective or subjective. Overall, the results suggest that descriptions fall midway on the continuum, with a slight bias towards the spectrum's thought-end, meaning that the recommendation to describe objectively, as proposed in the guidelines, is often not followed in practice. In terms of the strategies used, action descriptions most often use single verbs, whereas thought descriptions most often use verbs combined with an adverb or phrase. This could be explained by the different nature of actions and thoughts in visual communication. Since the former are concrete and can be presented directly, they can be described precisely without further qualification.

On the other hand, the latter are abstract and are always given indirectly, which may urge the describer to include some form of subjective qualification. However, the analysis showed that actions and thoughts could not be treated as entirely separate from each other: the characters' mental states are expressed mostly in descriptions that refer to their actions (81 percent). Also, descriptions of characters' actions are frequently accompanied by a reference to the emotion or reasoning behind that action (27 percent). We duly acknowledge that these results must be treated carefully and that the study presented here has clear limitations, which are predominantly due to the small scale of the corpus. Further research on a larger-scale corpus will have to be conducted to confirm or reject these findings. The same goes for the differences in thought descriptions that were observed when comparing the three episodes. We hypothesized that these are related to the differences in genre, but this will have to be studied in a much larger corpus. Our findings are also based on a Dutch-spoken corpus, and the results may be different for languages that use various syntactic and linguistic means to express actions and thoughts. Finally, one element that was not included in our analysis is which types of emotions are usually described. It is possible that universal emotions, such as anger or sadness, are easier to express visually. Therefore, it is easier to recognize and describe other emotions such as disgust or pride (Ekman and Friesen 2003; Ekman, Friesen and Ellsworth 1972; Faigin 2008). This may also result in different strategies being used to describe them and other subjectivity levels in the description.

To summarize, the differences between the visual and verbal modes of communication impact the practice of audio description. In this paper, we barely scratched the surface. It would be good to see future research uncover more of the implications of these differences on AD and other forms of intermodal translation.

References

AENOR 2005. *UNE 1530220 Audiodescripción para personas con discapacidad visual: Requisitos para la audiodescripción y elaboración de audioguías*. [Audio description for visually impaired people: Guidelines for audio description procedures and for the preparation of audio guides]. Madrid, AENOR.

Benecke, Bernd 2014. Character fixation and character description: The naming and describing of characters in Inglorious Basterds. In: Anna Maszerowska, Pilar Orero and Anna Matamala (eds.) *Audio description: New perspectives illustrated.* Amsterdam: John Benjamins, 141–157.

Benecke, Bernd and Elmar Dosch 2004. *Wenn aus Bildern Worte werden: Durch Audio-Description zum Hörfilm*. [When pictures become words: Making film accessible through audio description]. München: Baeyerischer Rundfunk.

Ekman, Paul and Wallace V. Friesen 2003. *Unmasking the face: A guide to recognizing emotions from facial expressions*. Cambridge, MA: Malor Books.

Ekman, Paul, Wallace V. Friesen and Phoebe Ellsworth 1972. *Emotion in the human face: Guidelines for research and an integration of findings*. New York, NY: Pergamon Press.

Emmott, Catherine 1997. *Narrative comprehension: A discourse perspective*. Oxford: Oxford University Press.

Faigin, Gary 2008. *The artist's complete guide to facial expression*. New York, NY: Watson-Guptill.

Fresno, Nazaret 2016. Carving characters in the mind: A theoretical approach to the reception of characters in audio described films. *Hermēneus* 18: 59–92.

Fresno, Nazaret, Judit Castellà and Olga Soler-Vilageliu 2016. 'What should I say?' Tentative criteria to prioritize information in the audio description of film characters. In: Anna Matamala and Pilar Orero (eds.) *Researching audio description*. 143–167. doi:10.1057/978-1-137-56917-2_8

Fryer, Louise 2016. *An introduction to audio description: A practical guide*. Abingdon: Routledge.

Georgakopoulou, Yota 2010. Audio description guidelines for Greek: A working document. In: Sonali Rai, Joan Greening and Leen Petré (eds.) *A comparative study of audio description guidelines prevalent in different countries*. London: Media and Culture Department, Royal National Institute of Blind People, 105–108.

Henkens, Mirjam 2014. *Fysieke kenmerken, gedrag en non-verbale communicatie van personages in audiodescriptie (AD).* [Physical characteristics, behavior, and non-verbal communication of characters in audio description (AD)]. Unpublished Master thesis, Department of Translation, University of Antwerp.

Jakobson, Roman 1959. On Linguistic Aspects of Translation. In: Reuben A. Brower (ed.) *On translation*. Cambridge: Harvard University Press. 232–239.

Punctum. International Journal of Semiotics | **06:01:**2020
ISSN 2459-2943 | DOI: 10.18680/hss.2020.0005 | **punctum.gr**

Kaindl, Klaus 2013. Multimodality and Translation. In: Carmen Millán and Francesca Bartrina (eds.) *The Routledge handbook of translation studies*. Abingdon: Routledge. 257-269.

Kress, Gunther R. 2003. *Literacy in the new media age.* Abingdon: Routledge.

Kress, Gunther R. and Theo Van Leeuwen 2001. *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.

Kruger, Jan-Louis 2010. Audio narration: Re-narrativising film. *Perspectives* 18 (3): 231–249. doi:10.1080/0907676x.2010.485686

Margolin, Uri 2007. Character. In: David Herman (ed.) *The Cambridge companion to narrative*. Cambridge: Cambridge University Press, 66–79.

Margolin, Uri 1990. Individuals in narrative worlds: An ontological perspective. *Poetics Today* 11 (4): 843–871. doi:10.2307/1773080

Monaco, James 2009. *How to read a film: Movies, media and beyond*. 4th ed. Oxford: Oxford University Press.

Morisset, Laure and Frédéric Gonant 2008. *L'audiodescription: Principes et orientations.* [Audio description: Principles and guidance]. Available here.

Ofcom 2017. *Code on television access services.* Available here.

Palmer, Alan 2004. *Fictional minds*. London: University of Nebraska Press.

Palmer, Alan and Andrew Salway 2015. Audio description on the thought-action continuum. *Style* 49 (2): 126–148. doi:10.5325/style.49.2.0126

Rai, Sonali, Joan Greening and Leen Petré 2010. *A comparative study of AD guidelines prevelant in different countries*. London: Media and Culture Department, Royal National Institute of Blind People.

Remael, Aline, Nina Reviers and Gert Vercauteren (eds.) 2015. *Pictures painted in words: ALAB audio description guidelines*. Trieste: EUT Edizioni Università di Trieste.

Rimmon-Kenan, Shlomith 2002. *Narrative fiction: Contemporary poetics*. London: Routledge.

Salway, Andrew 2007. A corpus based analysis of audio description. In: Jorge Diaz-Cintas, Pilar Orero and Aline Remael (eds.) *Media for all: Subtitling for the deaf, audio description and sign language*. Amsterdam: Rodopi, 151–174.

Schmid, Wolf 2014. The selection and concretization of elements in verbal and filmic narration. In: Jan Alber and Per K. Hansen (eds.) *Beyond classical narration: Transmedial and unnatural challenges*. Berlin: de Gruyter, 15–24.

Semino, Elena 2017. Corpus linguistics and metaphor. In: Barbara Dancygier (ed.) *The Cambridge handbook of cognitive linguistics*. Cambridge: Cambridge University Press, 463–476.

Snyder, Joel 2005. Audio description: The visual made verbal. *International congress series* 1282: 935–939. doi:10.1016/j.ics.2005.05.215

Snyder, Joel 2010. *Audio description guidelines and best practices*. Available here.

Stefanowitsch, Anatol 2007. Corpus-based approaches to metaphor and metonymy. In:

Anatol Stefanowitsch and Stefan T. Gries (eds.) *Corpus-based approaches to metaphor and metonymy*. Berlin: Mouton de Gruyter, 1–16

Vercauteren, Gert 2007. Towards a European guideline for audio description. In: Jorge Diaz-Cintas, Pilar Orero and Aline Remael (eds.) *Media for All: Subtitling for the deaf, audio description, and sign language*. Amsterdam: Rodopi, 139–149.

Vercauteren, Gert 2012. A narratological approach to content selection in audio description: towards a strategy for the description of narratological time. *MonTI: Monografías de Traducción e Interpretación* 4: 207–231. doi:10.6035/monti.2012.4.9

Vercauteren, Gert and Pilar Orero 2013. Describing facial expressions: Much more than meets the eye. *Quaderns: Revista de Traducció* 20: 187–199.

Walczak, Agnieszka and Louise Fryer 2017. Creative description: The impact of audio description style on presence in visually impaired audiences. *British Journal of Visual Impairment* 35 (1): 6–17. doi:10.1177/0264619616661603

AUTHOR

Bonnie Geerinck is a PhD student in Translation Studies at the Department of Applied Linguistics, University of Antwerp, Belgium.

Gert Vercauteren is Assistant Professor at the Department of Applied Linguistics, University of Antwerp, Belgium.

ISSN 2459-2943
http://punctum.gr

PUBLISHED BY
THE **HELLENIC**
SEMIOTIC SOCIETY