



LUND
UNIVERSITY

A Cross-Varietal Comparison of Audio Description:

Atypical (United States) vs. Totally Completely Fine (Australia)

Kristina Grinenko

Supervisors: Vasiliki Simaki, Jordan Zlatev
Centre for Language and Literature, Lund University

MA in Language and Linguistics, English Linguistics

SPVR01 Language and Linguistics: Degree Project – Master's (Two Years) Thesis, 30 credits

June 2024

Abstract

Audio Description (AD) is developing fast, becoming the norm for big streaming platforms. However, the research field remains largely underexplored. When comparing AD practices between countries, studies seem to focus on guidelines rather than how these are used, or not, by audio describers. Additionally, not enough emphasis is placed on cross-varietal differences of English AD, and in cases when such differences are considered, it is mostly the British and the US varieties that are juxtaposed. Very few have extended this comparison to other English-speaking countries, such as Australia.

Rather than looking into guidelines, this thesis asked what features and trends can be identified in AD when two AD tracks are analyzed. The features were systematized to make up five categories: Local coherence, Global coherence, Subjectivity/Objectivity, Formality and Character introduction. The study went on to ask how the AD tracks in the two shows can be compared in terms of each of these categories with their respective features.

The analysis revealed that, when compared to the US show *Atypical*, the Australian show *Totally Completely Fine* exhibits more explicit links within one scene and fewer cases of disturbed local coherence, as well as more specific cues signaling the start and end of flashbacks. It also has more subjective features such as interpretative verbs and named emotions, with fewer facial expressions described by referring to parts of the characters' face. Besides, the Australian show has fewer formal features such as Latinate verbs and complex grammar structures, but the average sentence length between the shows is similar. Finally, when characters are introduced on the Australian *Totally Completely Fine*, the audio describer exhibits a preference for their "permanent" features such as race and age, while their US counterpart also opts for the exclusively "temporary" traits like the clothing. Further research will show if these characteristics are typical for US and Australian traditions of AD more generally.

Keywords: Audio Description, cross-varietal study, intersemiotic translation, coherence, subjectivity, formality, character introduction

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Lund University for admitting me to the program and offering me a generous scholarship. This has been a life-changing opportunity for me, and I will always be grateful to the university and Sweden for this experience.

Secondly, I would like to express my sincere gratitude to my two supervisors, Jordan Zlatev and Vasiliki Simaki, who have helped me immensely on this long journey. Dear Vaso and Jordan, you have been incredibly supportive, and your feedback has always been constructive and encouraging. You both motivated me to move forward and believe in myself, while the high standards you set for your students helped me to do my best. Thanks to you, I can be proud of how my thesis turned out!

And lastly, I want to thank my classmates and friends, and especially Sarah Erasmus, for spending these months with me, encouraging me, saying I'm smart and my thesis is going well in the darkest of moments. A special thank you for helping me with English. You are a great friend and your thesis is going to be phenomenal!

Table of contents

CHAPTER 1. INTRODUCTION	6
CHAPTER 2. BACKGROUND.....	9
2.1 AD PRACTICE, AD TRACK AND AD TEXT.....	9
2.2 AUDIO DESCRIPTION: HISTORY AND APPROACHES TO STANDARDIZATION	11
2.3 AD AND INTERSEMIOTIC TRANSLATION	16
2.4 RE-CREATING SOURCE COHERENCE IN AD	19
2.5 NARRATIVE AND NARRATIVE ELEMENTS	21
2.6 SUBJECTIVITY AND OBJECTIVITY	23
2.7 LANGUAGE OF AUDIO DESCRIPTION: A DISTINCT GENRE?	25
2.8 HOW CHARACTERS ARE INTRODUCED	28
2.9 SPECIFIC RESEARCH QUESTIONS	30
CHAPTER 3. MATERIALS AND METHODS.....	31
3.1 MATERIALS.....	31
3.2 DATA COLLECTION	32
3.3 ANNOTATION AND COMPARISON SCHEME	33
3.4 PROCEDURE	35
3.4.1 <i>Category A: Local coherence</i>	35
3.4.2 <i>Category B: Global coherence (narrating flashbacks and montages)</i>	37
3.4.3 <i>Category C: Subjectivity/Objectivity</i>	37
3.4.4 <i>Category D: Formality</i>	38
3.4.5 <i>Category E: Character introduction</i>	41
CHAPTER 4. RESULTS.....	42
4.1 LOCAL COHERENCE.....	42
4.2 GLOBAL COHERENCE: NARRATING FLASHBACKS AND MONTAGES	42
4.3 SUBJECTIVITY/OBJECTIVITY	44
4.3.1 <i>Interpretative verbs (appear, struggle, try, attempt)</i>	44
4.3.2 <i>Named emotions</i>	45
4.3.3 <i>Described facial expressions</i>	45
4.4 FORMALITY	46
4.4.1 <i>The use of Latinate vs. Germanic verbs in the AD</i>	46
4.4.2 <i>Phrases with as and before, wh-clauses</i>	47
4.5 CHARACTER INTRODUCTION	48
CHAPTER 5. DISCUSSION	50
5.1 THE MAIN FEATURES OF AD	50
5.2 CURRENT TRENDS IN AD	50
5.2.1 <i>Some facial expressions are underdescribed</i>	50
5.2.2 <i>Using “frown” as a general term for various facial expressions</i>	53
5.2.3 <i>Using repetitions in AD</i>	55
5.3 LOCAL COHERENCE.....	56

5.4 GLOBAL COHERENCE	56
5.5 SUBJECTIVITY/OBJECTIVITY	57
5.6 FORMALITY	57
5.7 CHARACTER INTRODUCTION	58
CHAPTER 6. CONCLUSIONS	59
REFERENCES	63
APPENDIX 1: PLOT SUMMARIES	70
APPENDIX 2: AVERAGE SENTENCE LENGTH	71
APPENDIX 3: LOCAL COHERENCE	72
APPENDIX 4: FLASHBACKS	75
APPENDIX 5: REPETITION IN AD	78

CHAPTER 1. INTRODUCTION

Audio Description (AD) is the established English term for the institutionalized practice of producing speech on an additional audial track woven between dialogue lines into audiovisual content. AD is used for films, TV shows, musicals, theatrical performances, stand up shows and other content with the purpose of making it accessible to visually impaired audiences (Şulha, 2023). Providing AD makes it possible to include people with low or no vision into all kinds of public discussions, as well as giving them an opportunity to enjoy a wider variety of entertainment. The practice involves a team of professionals, often including the creative team of the show as well, with the goal of making the artistic product both accessible and enjoyable. With these two aspects in mind, researchers and practitioners remain committed to testing new approaches, comparing existing practices and gaining insights from the users themselves. Users, creators and researchers can make a meaningful contribution to a more inclusive world by constantly improving AD and demanding this service where it is not yet provided.

As a service, the AD practice was first launched in the United States in 1990, when 32 television stations started offering “a prerecorded narration of a TV program’s visual effects” (Gibson, 2021). After decades of advancement and despite consistent efforts, AD is still not provided in most parts of the world. When a country starts developing AD standards, it can model them after those already existing in other places and adjust to the needs of the specific language. However, since languages differ in how they encode objects and actions of the real world (e.g., Slobin, 1987), AD practices cannot be fully universal. For instance, *Netflix* AD guidelines recommend using precise verbs of manner, such as *hobble*, instead of a verb modified with other elements, e.g., *walk with difficulty* (*Audio Description Style Guide v2.5*, 2023). This would not be possible in many other languages, which usually do not express the categories of Motion and Manner in the verb (Bourne & Hurtado, 2007). By comparing how AD is realized in different languages, researchers gain a deeper understanding of what constitutes the language of AD, how diverse it can be as well as what approaches work best. Awareness of the differences between languages and cultures also contributes to making the practice of AD more widespread. It may be a more efficient option to translate AD from the original language, provided that translators are

aware of all the specifics of AD texts in both languages, since making an AD track from scratch for a translated film can be daunting for streaming platforms and producers.

To the extent that differences in AD practices across languages can be explained by typological dissimilarities and other language-specific norms (Arma, 2012), it could be expected that two AD tracks in the same language would be rather alike. However, this is not the case when US and British AD practices are compared. One notable difference between the two approaches is that the US AD is said to be more “objective” than the one across the Atlantic, according to *ITC Guidance On Standards for Audio Description* (2000: 15). For instance, when the British audio describer may say *She shoots him a disapproving look*, the US one might instead opt for *She raises an eyebrow and purses her lips*. Thus, the US standards require describing the visual cues and letting the audience draw their own conclusions instead of interpreting the actor’s facial expression. However, some researchers present *subjectivity* in AD as beneficial in specific cases (Soler Gallego, 2019). This controversial and much debated issue is discussed in Section 2.6.

While there are some (albeit, not many) comparative analyses between AD guidelines in different English-speaking countries, very few studies make detailed comparisons between different AD tracks (e.g., Diget, 2019) in order to gain a better understanding of the choices made by the audio describer. Sighted audiences receive many visual cues allowing them to build a coherent understanding of the multimodal and polysemiotic text. It is far from obvious which cues, from the multitudes accessible, played crucial roles in this process. The task of the audio describer is to select the most relevant visual representations, which will make building a coherent narrative effortless for the AD users. This challenging undertaking is additionally compounded by the notoriously tight time constraints (i.e., AD guidelines nearly universally prohibit overlaps with the dialogue). The links are created between different visual cues, but also between pieces of the dialogue, various sound effects, voices that need to be identified in group scenes and many other elements (Braun, 2011).

It is also relevant to compare AD tracks made in different varieties of the same language, given that, for example, as mentioned above, the US and British practices differ with respect to recommendations on subjectivity. This thesis deals with such a comparison, focusing on the AD tracks of two television shows: a US and an Australian one. This choice can be justified as follows. In the largely underexplored field, very little is known about the Australian AD practice, which emerged not so long ago (national broadcasters *ABC* and *SBS* started audio describing select

content in 2020, according to Ellis & Kent, n.d.). This thesis, therefore, juxtaposes the AD practices followed in two varieties of English: one relatively well studied variety (US) and one understudied variety (Australian). Given that there are so few comparative analyses, this thesis aims to develop a coding scheme where features and trends observable in the AD tracks under study can be identified and systematized. This scheme can serve as a contribution to further research comparing AD tracks.

The two overarching research questions to be elaborated on at the end of Chapter 2 are the following:

RQ 1: What are the main features and current trends in Audio Description?

RQ 2: What are the commonalities and discrepancies between AD tracks in different varieties of English?

In addressing these research questions, the study will investigate different aspects of AD, such as the development of AD around the world, an AD-specific language, subjectivity, coherence and other relevant concepts. The methods chapter proposes a coding scheme for comparing AD tracks, which outlines the categories as well as the annotation process. The remainder of this thesis is organized as follows: Chapter 2 presents the background of AD as a practice and a field of study, Chapter 3 introduces the methods used in the research and describes the data as well as the ways in which it was collected and organized. These are followed by Chapter 4, which delves into the findings, and Chapter 5, where the results of the study and their implications are discussed. The conclusions are presented in Chapter 6, along with the possible venues for future research.

CHAPTER 2. BACKGROUND

2.1 AD practice, AD track and AD text

As stated in Chapter 1, Audio Description (AD) is above all a practice of making audiovisual content accessible for blind and partially sighted audiences by inserting voiceover descriptions and narrations (see Section 2.5 below) of salient visual elements. But here I wish to introduce two concepts that “AD” is sometimes used for as well, which are only parts of AD practice: the AD track and the entire AD text.

To get an idea of how these voiceover descriptions work, consider Figure 1. Two audio tracks for the same language are shown on top of each other: the upper track is the original audio of a film with the dialogue, background noises, narrator (if any) and music. The lower track has all the above-mentioned features plus AD insertions during the dialogue pauses.

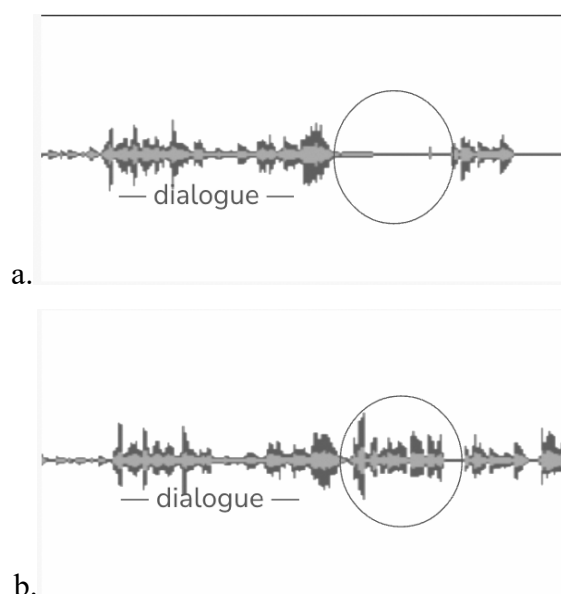


Figure 1: (a) The original audio track and (b) AD track, taken from *Atypical* (Netflix), season 1, episode 1

An example of such AD insertions is given in Table 1 below. This is the first episode of the show *Totally Completely Fine*, where the name of one of the characters is given for the first time. Since it is a scene with several people, the sighted audience knows who is being addressed because that character (Hendrix) is in the frame. The AD track compensates for this by alluding to a physical

trait (mustache), which was mentioned by the audio describer earlier. The audio describer thus managed to establish links between different scenes allowing the visually impaired audience to form a coherent understanding.

Table 1: AD from Totally Completely Fine; season 1, episode 1

01x01	02:14	AD	She passes a burly white man with a <u>mustache</u> on the phone.

	05:55	Dialogue	Lawyer: Now, Hendrix!
	05:56	AD	The man with the <u>mustache</u> .

It is typically recommended to avoid an overlap of AD and music (Braun, 2021), but in reality, this is rarely possible. Indeed, the characters tend to do something that needs to be described while the music is playing, so musical insertions are treated as an opportunity to fit in more details. If pauses in dialogue contain music, its volume is usually decreased so that the AD track may come to the foreground; this is shown in the higher amplitude of that part in Figure 2.

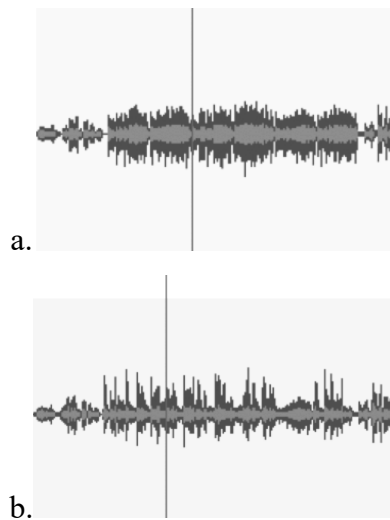


Figure 2: (a) Original track with music and (b) AD track with music and audio description of Atypical (Netflix); season 1, episode 1

Interestingly, it appears that the AD track is not simply added to fill the pauses, but instead, the original track gets remixed. Figure 3 shows the original track (a) and AD track (b) aligned at the beginning of the section. In *b* (AD track), a short delay is seen in the middle of the excerpt, which becomes longer towards the end. This indicates that some of the other elements, such as the onset of dialogue and narrator’s voice have been pushed back to make more space for the AD track.

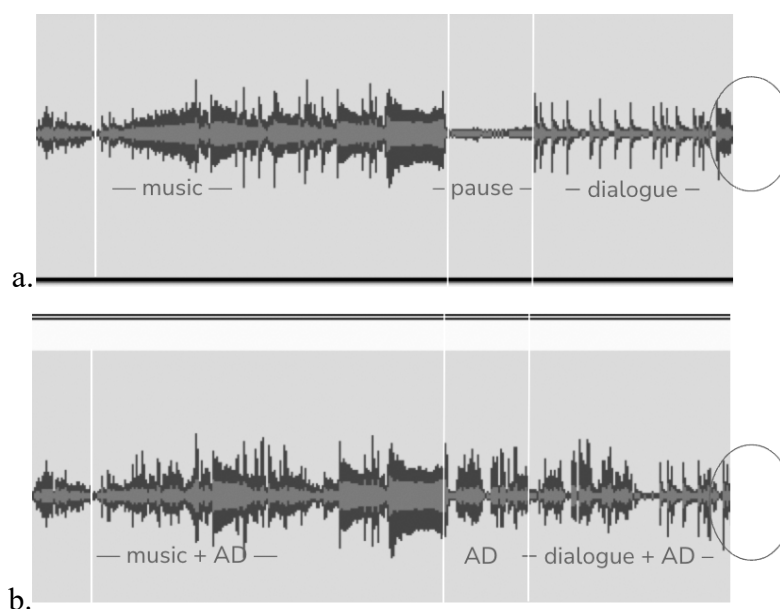


Figure 3: Original audio (a) and AD (b) aligned, white lines and circles added to show a shift in the AD track, *Atypical* (Netflix); season 1, episode 1

These are some of the examples showing that an AD track is not only about insertions. Rather, it is a whole new track with remixed music and dialogue, where all the elements are intertwined. The whole *AD text* is created by a team of professionals, who write the script, record it and remix the music and the dialogue to make the content of the text as accessible as possible for a new audience.

2.2 Audio Description: History and approaches to standardization

The importance of the practice of AD is becoming clear in many places of the world, with unique approaches adopted and very different terms used, such as *syntolkning*, “vision interpretation” in Sweden, *tyflokomentuvannya*, “blind commenting” (from the Greek *tiflós*, “blind”) in Ukraine, etc. In this text, I use AD as the term for the AD practice in general, irrespective of such, perhaps not insignificant, terminological differences.

The adoption and standardization of AD practices have been a slow process. According to the UK's communications regulator *Ofcom* (2006), the first ever audio described product was made by the Japanese broadcaster *NTV* in 1983. Although presently Japan is not among the leaders in AD and, as of 2016, did not have official guidelines, the concept of a Japanese third-party film narrator, *benshi*, may have contributed to this early start (Sirés, 2016). While *benshi* (lit. “orator”) are still popular today, they were especially in demand in the era of silent films and were hired by movie theaters to describe what happened on the screen.

Germany published the first AD guidelines created by Dosch and Benecke in 1997: *Wenn aus Bildern Worte Werden. Ein Handbuch für Filmbeschreibung* (“When Images Become Words. Handbook for Film Description”). They served as a blueprint for other guidelines released in Europe in the first decade of the new millennium. The main principles have not changed to this day, with rules covering content selection or “what to describe”, AD style or “how to describe” and timing or “when to describe” (Tor-Carroggio & Vercauteren, 2020b).

Today, official AD guidelines are available in most European countries, the US, Canada and Australia, with each new country often building upon the rules formulated by their predecessors. This normalization of the AD practice naturally leads to a lot of similarities across the countries, but there are still discrepancies. For instance, the question of when to first mention the name of a character is approached differently in different official guidelines. The German guidelines maintain that the character should be named only when someone calls them by their name or it becomes known otherwise (*Audio Description Coalition* 2009, p. 7), while the UK's *Ofcom* (ITC 2000, pp. 16–17) recommends giving the name as early as possible, even if it was not mentioned in the film. Approaches to style also vary, notably between the prominent AD practices in the US, the UK and Spain¹. The approaches to style between Spanish and English differ significantly (see Section 2.7 for the language of AD). The level of detail can also be remarkably dissimilar between the languages, which can be seen from a comparison of the AD tracks made independently in the two languages for the same film *The Hours* (Bourne & Hurtado, 2007), as shown in (1) and (2).

¹ These three countries are considered leaders of the field based on the year of AD inception, considerable advancements in the offer of audio described television, cinema or theater, in national legislation, oversight bodies and academic interest (Franca, 2016).

(1) *The woman's body, face down, is carried by the swift current through swaying reeds along the murky river bed, her gold wedding band glinting on her finger, a shoe slipping off her foot.*

(2) *El cuerpo sumergido de Virginia es arrastrado por la corriente.*

[‘The submerged body of Virginia is swept away by the current.’]

As mentioned in Chapter 1, such disparities in the level of detail may also stem from grammatical and pragmatic distinctions between the different languages themselves (e.g., Slobin, 1987). Presenting the comparison of these contrasting takes on *The Hours*, Bourne and Hurtado (2007) highlight some issues that should be considered if an AD track were to be translated from one language to another, rather than made from scratch in each new language. Provided that the translator is aware of the specifics of both the source and the target language when it comes to the language rules and the practices of AD, such a translation can be a step in promoting accessibility. Moreover, this can lay the groundwork for standardizing AD in countries where it is not so prevalent at the moment. But this needs to be done with care, due to cross-cultural and cross-linguistic differences.

Explaining the challenges of creating AD for Arabic, H. S. Alattar (2021) emphasizes the need for more research in the area. Currently, many countries in the Middle East and Africa cannot produce AD because of significant expenses the workflow entails. Apart from translation, a possible solution to this problem could be to use synthetic speech, as done by the Russian company *Rufilms*²; also see Szarkowska (2011). Text-to-speech technologies are constantly improving, and this would help significantly minimize the production costs for AD. This, however, implies a risk for a “dehumanization” of the practice, as it may lead to a loss of nuanced interpretation that a human audio describer can convey through voice. Finding a balance between affordable accessibility in more countries and the luxury of seeing the production of the AD text as an art form seems to be required.

With the goal to make AD available for the general public, more and more countries offer a mobile application that automatically synchronizes the film playing in a movie theater or on TV at home with an AD track available in its database, making more content accessible for visually impaired people. The app *MovieReading* was initially launched in Italy and spread to other countries, such as Sweden (Redden, 2015) and Australia (Cussen, 2013). In South Africa, a similar

² <https://rusubtitles.com/tiflo/>

app was launched by Shakila Maharaj (Chemaly, 2021), who also encourages users to contribute and add their own AD tracks to the database³. This is an interesting development in a different direction from some other examples provided earlier, involving more personal engagement and less standardization.

A noteworthy case when it comes to AD is China, since the practice there developed separately from the West (Tor-Carroggio & Vercauteren, 2020). Like in Japan and a number of other countries, there are no official AD guidelines here, but the China Braille Library (Beijing) audio-describes approximately 50 movies every year and records the AD of 20–30 movies on CD. Despite having been developed independently, their guidelines have many similarities with those issued in other places with regard to content selection and style. A recommendation that stands out states that objects should be described based on a tactile perspective, e.g. *soft as a pillow* or *thin as a sheet of paper*. However, the main difference between the European and Chinese AD standards concerns the attitude to *subjectivity* (see Section 2.6). According to Tor-Carroggio & Vercauteren (2020), Europe is moving away from the dichotomy between objectivity and subjectivity, while any perspective-taking in the Chinese guidelines is explicitly and strongly discouraged (“language needs to be objective, concise, accurate and formal”).

A less prescriptive approach to AD is presented by Fryer (2018) in a paper entitled *The independent audio describer is dead: Long live audio description!* Focusing on AD in theater, she maintains that AD as we know it (Traditional AD, or TAD) is outdated and the future lies in the so-called Integrated Audio Description (IAD). She criticizes the old approach, which she calls “ocularcentric”, meaning that the information is presented from a sighted viewpoint.⁴ Fryer (2018) characterizes TAD as exclusive, neutral, non-auteur, third-party and post hoc. Each of these characteristics has a counterpart in its supposed successor, IAD. Namely, instead of separating the audience into blind and sighted groups (so the sighted are “not bothered”, as the creators of the AD app say), IAD is made to be part of the experience for everyone to enjoy⁵. Fryer connects inclusivity

³ Maharaj became blind in her 20s and had a strong visual memory, but she could not enjoy films fully because of her deteriorating vision. Learning about the best practices of AD in other countries, she was able to start her own company and cooperate with local filmmakers.

⁴ It could be argued, however, that AD is usually reviewed and edited by a blind person precisely to avoid this problem (as stated on the *Rufilms* website).

⁵ An example of such a turn in accessibility is the ASL interpreter Justina Miles, who became a star after performing with the singer Rihanna at the American Super Bowl championship (O’Kane, 2023). This proved to be unifying and inspiring to many and contributed to a positive representation for the deaf community.

with the principles of Universal Design (UD) — a product should “be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design” (Connell et al., 1997). Therefore, IAD is recommended to be produced in an entirely different way from the (so-called) TAD. Instead of an independent describer, as the title of the paper suggests, a new collaborative process between the AD team and the production team is called for.

Furthermore, the idea of TAD being “neutral” and “third-party” has been problematized based on the following observation: *The ADLAB Guidelines* (Remael, Reviere & Vercauteren, 2015) urge the describer to “determine” 89 times, “decide” 102 times and “choose” 7 times when creating the script⁶. This independent decision-making inevitably results in interpretation by the third-party and may raise doubts in the minds of AD users as to the reliability of such information. With IAD, on the other hand, all decisions are made in cooperation with the production team, including the director, the artists, and others. Instead of being post hoc and having to be adjusted to the content, it is written, rehearsed and recorded with the whole artistic team from the very beginning, making the product more inclusive than ever. In this way, the AD track becomes part of the final product on the same level as other elements. One example of such an adjustment on the part of the team comes from Nathan Geering, owner of a hip hop dance company *Rationale*, who says that he sometimes extends a sequence of dance moves to give the describer enough time to explain what is happening (Fryer, 2018).

Audio Description is, as discussed in the following section, a form of translation practice – and as such, it is not the first that starts to be approached creatively once standardized enough. A comparison can be drawn with *transcreation*, which is a type of textual work almost as common as others, such as legal translation or editing. Unlike the latter, however, transcreation implies that the person working on the text (usually a marketing copy) acts as its co-author in the new language. A transcreator makes a new copy with references and selling points familiar to their local audience, while preserving the main idea. Although a recognized practice today, it was initially regarded as an unnecessary task involving too much freedom on the part of the translator (Gaballos, 2012).

Another example from the field is translation of poems, which is known to be one of the most challenging forms of translation (Jakobson, 1959). The British translator and poet Sasha Dugdale who works with the Russian poet Maria Stepanova shares that their approach is more about co-creation than translation. She says, “Maria gave me a lot of permission to translate with

⁶ The figures refer to the number of times these verbs are used in the guidelines.

freedom, and that was incredibly exciting.” The person leading the talk about the translation process remarks, “In many ways I feel that with your translations of Maria you are *continuing* her poems in English, you are *adding* to [them]”, to which Dugdale comments, “I feel that... the basis of the translations is in the many conversations we’ve had. And talking about all sorts of things, really... [Maria’s voice in my head] transfers into [my] literary translations...” (Bloodaxe Books, 2021). This approach resembles some of the aspects of IAD, when the translator uses the background context, rather than only the information explicitly given in the source, and even has access to the author of the original text. This discussion becomes even more relevant when it comes to translation between two semiotic systems, which is covered in the next section.

2.3 AD and intersemiotic translation

Jakobson (1959) was the first to define the concept of intersemiotic translation, delineating between the translation of a message into different signs within the same language (intralingual), into a new language (interlingual), or into another semiotic system altogether (intersemiotic). The source system remained the same, language, but Jakobson’s definition was subsequently extended by other scholars, such as Nida, for whom intersemiotic translation meant “transference of a message from one kind of symbolic system to another” (1964: 4). A symbolic, or rather a semiotic system, as defined by Zlatev et al. (2023: 2), “consists of all the signs/signals of a particular type, and their interrelations”. Signs, like words, most gestures and pictures, are distinguished from signals like spontaneous postures and facial expressions, by having denotational (referential) meaning, which is differentiated from their corresponding expressions (Zlatev, Zywickzynski & Waciewicz, 2020). On this basis, Zlatev et al. (2023) define three universal human (macro) sign systems: Language, Gesture and Depiction, which can be realized in many different forms of media: speech, writing, film, sand drawing, etc.

Within the context of AD, Diget (2019: 16) defines intersemiotic translation as “translation from one act of communication and its employed semiotic systems to another act of communication that entails different, more or fewer semiotic systems.” As in all kinds of translation, this process implies a *double act of communication* (Sonesson, 2014), where the translator acts as an interpreter of the source message and, at the same time, as a creator of the target message, aiming at making the content and style of these overlap as much as possible. This does not really apply to many intermedial transfer processes, such as a painting based on a mythological scene or a book

adaptation to a film, since in these contexts, it is neither possible nor desirable to maintain the content of the source text.

However, AD does qualify as intersemiotic translation, since the audio describer aims to convey the message as close to the source as possible, rather than reimagine it or draw inspiration from it (as in the case with many other transfers). Achieving this closeness can be a particularly challenging task when the source and the target are expressed in different semiotic systems, because of inherent differences in meaning-making between their types of signs and signals (Zlatev et al. 2023).

Restrictions when selecting content to describe may be caused by the lack of appropriate vocabulary, among other things. After all, finding “vivid and precise” (as recommended, e.g., in Arma, 2011) words to describe things usually perceived through vision requires great literary mastery. The audio describer acts as both an interpreter of the visual message and a creator of its auditory equivalent while trying to remain as faithful to the source as possible (as all AD guidelines require), and thus performs intersemiotic translation in its purest sense. This is the case even if other terms are used, like “intersemiotic compensation” by Pujol (2007), who argued that the outcome can only *aspire* to correspond to the original content, while *full* equivalence is hardly achievable (“trying to compensate with words the aesthetic pleasure caused by images”).

However, even classical inter-lingual translation cannot exhibit such “formal” equivalence, as there are always differences between languages in how objects and events are categorized or construed. Rather, Nida and Taber (1982) made famous the notion of “dynamic equivalence”, stating that the goal in translation is to achieve comparable “effects” between that of the source content and that of the target content on their respective audiences. A study conducted by Uulu et al. (2021) revealed that visually impaired participants were capable of understanding and recounting the events to the same degree as sighted participants when the film included an AD track, and this could be understood as an example of achieving such dynamic equivalence.

An important aspect of any kind of translation is the translator’s ability, as both the receiver and the sender, to ensure the information is coherent within a larger context. In intersemiotic translation involving film, this means having background knowledge of common devices used in image construction (McGonigle, 2013), an example of which can be found in the AD track of *Black Swan* (2010), see (3) below. In this case, McGonigle points out that the *Black Swan* AD fails to mention three full-length mirrors in the room, adding that the use of mirror imagery is “indexical

of duality and a fractured self-image...”, which is relevant to how the character is presented in the story.


(3) *In the living room, she stretches out hands and one foot on the floor, the other straight up in the air.*

Baldry and Thibault (2005: 5) also highlight the significance of intertextuality, stressing its role in situating film texts within a network of interconnected constructions and messages. Semiotic elements within them are said to be “codeployed in ways that belong to a common intertextual pattern.”

As a type of intersemiotic translation, AD is performed by conveying in three semiotic systems (language, music, sound cues) and one sensory modality (auditory) what is conveyed in the source using five semiotic systems (depiction, language, gesture, music and sound cues) and in two modalities (auditory and visual), see Table 2 (Diget, 2019). On top of that, language is also coded in two modalities in the source text, since, apart from the dialogue and narrator’s voice, there is on-screen text, such as signs, emails the characters receive, etc., as well as credits and captions, e.g., *TWO WEEKS LATER*. As for audio cues, while sounds such as a knock on a door or tap water running may need no explanation, many others can be ambiguous. In those cases, the audio describer has to provide an intersemiotic translation of the visual support available to the sighted audience (Ofcom, 2001).

Table 2: Intersemiotic translation of semiotic systems and modalities of a film into those of an AD track (adapted from Diget, 2010)

Intersemiotic translation



CONTENT	SEMIOTIC SYSTEMS	MODALITIES
Film	Depiction, language, gesture, music and sound cues	Auditory and visual
AD	Language, music, sound cues	Auditory

The audio describer performing intersemiotic translation has to be aware of this, as well as other meaningful cinematic devices, since audiovisual content can be a complex artistic product with multiple layers of meaning (Diget, 2010). For this reason, in order to deliver a dynamically

equivalent translation, McGonigle (2013) points out that content selection has to be done according to the product genre. One example she gives is the AD in the film *Casino Royale* (2006), where the audio describer chooses to prioritize one signature element of the Bond series over another. Namely, the film ends with an animation sequence which remained beyond the scope of the AD because the music playing in the background was seemingly considered a more meaningful element for the genre (the Bond music is a big part of the producers' marketing campaign; Bellringer, 2015). In certain cases, McGonigle (2013) recommends forming neologisms for different kinds of montage and other technical terms such as *the scene fades, cut to, shot, frame*, etc., to make providing full information about an audiovisual product a standard approach in AD practice. This point resonates with Pujol's (2007) view of AD as "intersemiotic compensation", highlighting that still more things can be improved to make audiovisual content truly accessible.

Apart from all these considerations, there is the challenge of creating a coherent narrative structure by establishing links between different elements of the audiovisual product, which is the subject covered in the following section.

2.4 Re-creating source coherence in AD

Since the verbal AD track is the product of an intersemiotic translation of multimodal (audiovisual) and polysemiotic (expressed in images, inter alia) content, the target AD text needs to recreate the continuum of time, place and events, as pointed out by Braun (2011). As in any kind of discourse, the audience needs to be able to "build a coherent picture⁷ of the series of events being described and [to] fit the events together" (Brown and Yule 1983: 197).

Van Dijk distinguishes between *local* and *global* coherence (1977), where the former is created between conjoined utterances and the latter emerges from an overall discourse topic and consistency of elements such as register, style and choice of expressions. By drawing attention to the specific place of AD texts with regard to their coherence and cohesiveness, Braun (2021) adapts van Dijk's terms to AD: local coherence is built within individual scenes, while global coherence reaches out across scenes.

Ensuring connectivity in AD is challenging because of the inherent differences between the source and the target texts, expressed as explained in Section 2.3. In the depiction system, which is communicated in the visual modality, all items seen together are assumed to have some kind of

⁷Note that the authors are using the term "picture" in a metaphorical sense here.

connection, however the symbolic (i.e., convention-based) and sequential nature of language necessitates more explicit links between these items (*cf.* Braun 2011).

When it comes to sound-image connection, Braun (2011) provides several examples from the AD track of the film *Girl with a Pearl Earring*. She emphasizes that even for real-life sounds which, in principle, can be easily identified by visually impaired audiences (e.g., rattle of keys, chirping of a bird, etc.), coherence should be re-created by purposefully arranging the sentences in a way that ensures the sound is preceded by the mention of its source, as in (4). Braun further suggests another way of connecting image and sound, which entails even more explicit links (5).

(4) *Caterina, who is pregnant, carries keys. [KEYS RATTLE]... They pass a parrot on a perch. [A BIRD CHIRPS].*

(5) *Caterina, who is pregnant, carries keys, which rattle [KEYS RATTLE]... They pass a parrot that chirps. [A BIRD CHIRPS].*

A connection as straightforward as in (5) may be required for more ambiguous or unexpected sounds and, in general, there is no universally accepted approach to building continuity. Some guidelines even state that the sound may be described *after* the event, as it can be “more effective” this way (Snyder, 2010:28). Because of the complexity involved in creating audio descriptions, instances of “disturbed” coherence, as described by Braun (2011), are quite common. This is particularly true given the time constraints inherent to audio description.

Apart from sound-image connection, local coherence can be about the association between the AD and the dialogue. For example, audio describers often refer back to characters or items using the respective pronouns. An example shown in Table 3 is from *Totally Completely Fine*, referred to above.

Table 3: Local coherence in Totally Completely Fine

01x02	46:01	AD	Vivian checks the caller and answers it.
	46:10	Dialogue	John: It's... It's John.
	46:11	AD	He crouches by the sink in his bathroom.

The pronoun *he* refers to John, who appears in the dialogue but not in the previous part of the AD track. In this way, the AD text is built around the other elements, connecting them and entering in relationships with them.

Global coherence relies on the same principle, but the links are to be re-created between different scenes and even episodes of a show. Braun (2021) observes that the audio describer of the film *The Hours* used similar vocabulary when relaying similar gestures. The gestures are made by different characters and in different scenes, but they are meant to show a connection across time, which was successfully re-created for the visually impaired audiences.

Global coherence is especially important when it comes to describing flashbacks and scenes in characters' imagination. Since flashbacks often disturb coherence, even sighted audiences need some sort of assistance to immediately recognize a shift. For instance, a flashback would typically have a different light and the characters may be presented as younger versions of themselves, whether by changing their hairstyles or employing different actors. The transition between the present moment and the flashback may take the form of a flash, a blurred image, etc. These effects can help the sighted viewers to see that they are now in a flashback without the characters saying any specific cues such as "I remember...". The task of the audio describer is to make the start and the end of the flashback just as recognizable without overwhelming the audience with details such as the description of a different light, character clothes, etc. Kruger (2010) calls for a special attention to flashbacks as a "complex narrative element" and advocates for the term *audio narration* instead of *audio description*, as also discussed in the following section. From his standpoint, rather than simply substituting the visually expressed meanings with verbally expressed ones, audio narration "should consciously and consistently create a narrative text that will be accessible to the audience who does not have access to those visual codes that allow the sighted audience to activate the audio(+)visual narrative text." (ibid: 2010). This leads naturally to the following section, which explores the notions of narrative and narrative elements in relation to AD.

2.5 Narrative and narrative elements

In agreement with Kruger (2010), Vercauteren (2016) as well as studies in narratology (e.g., Bal, 1997) and cognitive semiotics (e.g., Stampoulidis, 2019), Diget (2019) proposes a three-level concept of narrative applicable to AD, as shown in Figure 4. On the surface level of a narrative, there is *narration* which is perceivable directly and is expressed in one or more semiotic systems.

On the “bottom” is the *fabula*: the chronological sequence of events that is being represented. In the middle is the *organization*, or the story itself – the manner in which the fabula is being construed, for example starting from the final event, and ending with the first one, as in the film *Memento* (Nolan, 2000). In a film, as well as in literature, this level refers to telling or showing the events of the fabula by arranging them in a certain way.

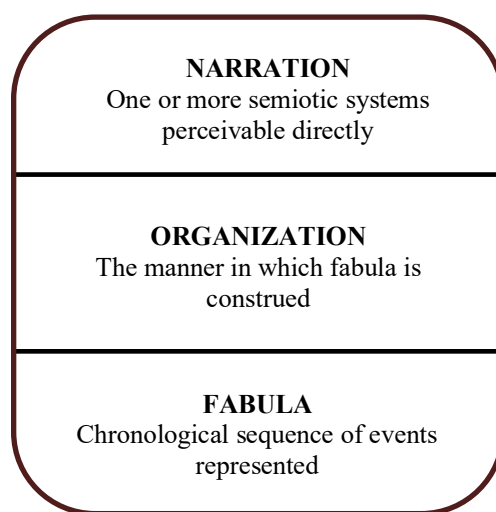


Figure 4: The three levels of narrative (Diget 2019)

Diget (2019) identified five categories as especially salient in the construction of narrative, called “the Five Ws”: (i) *where* (spatiality), (ii) *when* (temporality), (iii) *who* (characters), (iv) *what* (events), and (v) *why* (causal and other relations). Her comparative study of three AD tracks performed by three different audio describers of the same film found that more narrative elements corresponding to the Five Ws were conveyed visually than in the AD tracks: while these elements were identified in about 70% of the scenes in the visual form, they were narrated in only 41% of the scenes.

Arma (2011) argues that it is important to add the *how* category to the list, since visually impaired people must have access to the manner in which the actions are performed. This is the way the audience receives a significant portion of information about the characters’ personalities and the overall plot. This poses a challenge for languages, which do not express the categories of *what* and *how* in one word, like it is done in English (see Slobin, 2004). AD in these languages (e.g., Spanish and French) would often have to fill this category separately, e.g., using adverbs or

entire phrases. Because of time constraints, in many cases, they could possibly choose not to include the *how* category at all (see the example of Spanish in Bourne & Hurtado, 2007).

The *why* category can also be problematic because information about the reason or the cause of something can be encoded in ways other than just through visually presented details. It means that, in principle, this information is retrievable from the overall context both for the sighted audiences and for the visually impaired ones. So, if the audio describer decides to provide the reason for or the cause of what is happening (e.g., *She looks around to see if anyone is watching*), they take on the interpretation task that was meant for the audience, which could be regarded as unnecessary or condescending. This is the perspective adopted by The American Council of the Blind (2010, p. 17), among others. The opposite approach is proposed by Kruger, who maintains that in order to construct a coherent narrative, the *why* category (or “SO WHAT?” in addition to “WHAT” and “HOW”) is essential (Kruger, 2010). From his standpoint, if the audio describer wants to achieve the “narrative effect” and properly “re-narrativise” the source product, interpretation is unavoidable. The following section considers the matter of subjectivity and objectivity in AD.

2.6 Subjectivity and Objectivity

Subjectivity in audio description is a recurring theme in scholarly discourse. Advocating for the What-You-See-Is-What-You-Get principle, Snyder (2007), in particular, maintains that AD should be limited to observable details to avoid subjective interpretations. Similarly, Pfanstiehl and Pfanstiehl (quoted after Udo and Fels 2009: 179, as quoted in *The Play's the Thing* 1985: 91) liken describers to impartial lenses, cautioning against personal evaluations or assumptions in their descriptions. However, achieving such objectivity poses challenges. Udo and Fels (2009) argue that complete objectivity is unattainable, criticizing Snyder's approach. Hyks (quoted after Mazur & Chmiel, 2012) echoes this sentiment, maintaining that describers will have diverse interpretations due to individual perspectives and cultural influences. The impact of individual perception on describers' choices is also stressed by Pujol and Orero (2007), who highlight the inherent subjectivity in AD.

Identifying and analyzing subjectivity can be a challenging task, since it is so ingrained in us as language users and can often remain unnoticed. Moreover, a study by Mazur and Chmiel (2012) involving speakers of 11 languages, found that there is no agreement on subjectivity among

them. The participants were to watch a short film and recount the sequence of events, one of which was “Giving pears to boys”. Instead of saying that the pears were “given”, 82.3% of the UK participants interpreted the event as a gesture of thanks, while this figure only came at 38.3% among their Italian counterparts. What is more, 10% of the Spanish participants added a moral evaluation of the event (“as he should do...”), while most other groups refrained from giving a judgement entirely. The authors conclude that creating common European AD guidelines would be difficult given the many dissimilarities across the languages. It is proposed, therefore, to view subjectivity as a scale rather than a binary opposition, suggesting that one interpretation can be more subjective than the other (Mazur & Chmiel, 2012). This continuum is visualized in Figure 5.

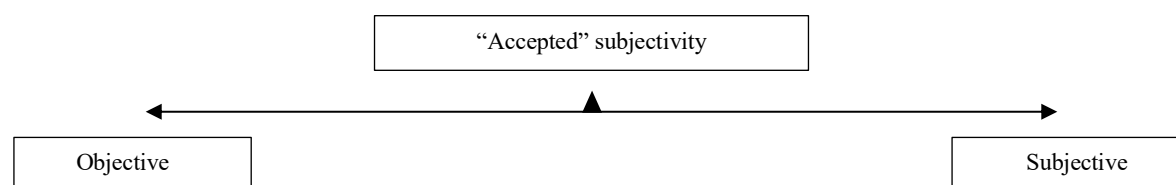


Figure 5: The subjectivity continuum in AD

Some AD choices may be considered overly interpretative. For instance, Remael (2015) warns against saying “She is having a nightmare” if a character is shown tossing around in her bed while sleeping. On the other hand, certain subjective descriptions are considered acceptable, especially the so-called “basic emotions”, such as “surprised” or “angry” (Vercauteren & Orero, 2013). Vercauteren and Orero argue that longer and more objective descriptions of these expressions would likely confuse the viewer. There are other reasons why audio describers may opt for a more (but to a generally “acceptable” degree) subjective approach. Namely:

- “A raised eyebrow” may point to a variety of emotions, from surprise to anger, which may lead to emotions being ambiguous and interchangeable (Dosch & Benecke, 2004, p. 24)
- Facial expressions described from the objective end of the continuum may be difficult to interpret for those AD users who have never been sighted and therefore may not know the entire range of “pursed lips” or other expressions (Mazur & Chmiel, 2012)
- Objective descriptions tend to be longer and harder to follow, as observed in an experiment with a focus group (RNIB & VocalEyes, 2003)
- Given time constraints, it is not always possible to include an objective description, in which

case the audio describer that wishes to adhere to objectivity requirements would likely choose to not give any information about the characters' expressions. In this situation, the viewers would probably benefit more from a short subjective description in which the emotion would be named in one word, rather than not have any insight whatsoever (Vercauteren & Orero, 2013).

Thus, when dealing with a facial expression, audio describers opt for one of the following approaches: describing by referring to parts of the character's face (objective extreme), naming the emotion (subjective extreme with exceptions for "basic emotions" and, perhaps, other cases) and omitting the emotion (outside the continuum; alternatively, could be interpreted as subjective).

Apart from the naming of emotions, researchers analyzing AD from this standpoint identified different elements that may be seen as a sign of subjective language. For instance, in Diget's research referenced earlier (2019), she identified instances of subjectivity in audio description where the describer interpreted a character's behavior, facial expressions, or body language as conveying a particular purpose or emotion. Categorizing subjectivity in visual art AD, Soler Gallego (2019) utilized Bednarek's classification of "opinion lexis" with dimensions such as expectedness, emotivity, importance, authenticity, power, reliability, and causality. In other studies, Soler Gallego categorized as forms of subjectivity evaluative adjectives, creative descriptions, emotional language (2023), and deliberate metaphors (Luque & Gallego, 2019). These studies have deepened our understanding of subjective language in audio description. The next section will look into other features commonly found in the language of AD.

2.7 Language of Audio Description: A distinct genre?

There is ongoing debate regarding whether the language used in AD qualifies as a specialized genre (e.g., Piety, 2004; Bourne & Hurtado, 2007). The guidelines are not especially helpful in determining what exactly this language form is, saying only that it has to be rich, precise and succinct (Audio Description Coalition, 2009). According to Arma (2011), who analyzed a large corpus of British AD, the language of AD tends to be "highly descriptive" and any sort of vagueness or ambiguity is rare.

It should be noted that the AD language is not homogenous, as pointed out in Section 2.2. It varies depending on the content (e.g., children's programs require age-appropriate vocabulary)

as well as across countries. For instance, when compiling the *TIWO* corpus of 91 films, Salway (2004) made sure to only include the British AD. However, the exact differences are yet to be explored and more corpus studies (including variety-specific ones) are called for.

Piety (2004) was the first to suggest that AD language was a distinct “language for special purposes”. These purposes entail providing information about events in cause-effect relationships occurring in space and time, as well as about the characters involved in the events and their emotional states. Therefore, this special genre is “shaped by communicative needs placed on audio description itself” (Arma, 2011).

Apart from, for example, the manner verbs in English, discussed in sections 2.5 and 2.6, a signature feature of the AD language would be the use of technical vocabulary or film jargon, such as the mention of how the light falls, as well as the description of montages, flashbacks and camera movements, as recommended by McGonigle (2013).

When it comes to sentence structure, the guidelines are in consensus prescribing short sentences over long ones (*Audio Description Coalition* 2009: 6). The tense in English is universally simple present, with rare inclusions of present continuous and present perfect. ITC (2000) suggests an occasional use of the present participle (gerund), to give “a better narrative feel”, since “[i]f the simple present is used throughout, it can sound abrupt. Where there is the luxury of enough time, a description should read like a piece of writing that makes sense on its own.” However, it is often recommended to choose the present participle as part of a non-finite phrase rather than in a continuous form, e.g., *Stomping up the stairs, he...* will be preferred to *He is stomping up the stairs* (Arma 2011). Overall, Arma notes that gerund is a good time-saving solution because it helps to avoid relative clauses, which are longer.

Since saving time is one of the most important considerations for audio describers, many grammatical choices seem to be linked to it. For instance, Bourne and Hurtado highlight connectors such as *while* and *as*, which are remarkably frequent in the English AD of the film *The Hours* (2007). Whereas these two conjunctions are used to connect two different agents in a sentence, describing two actions by one character involves another typical structure: the present participle phrase mentioned above (*Stomping up the stairs, he...*).

Drawing on Halliday’s discussion on spoken and written language (Halliday, 1978), Arma places the language of audio description on the spoken/written continuum, stating that it has characteristics of both (Arma, 2011). One of the spoken-language features she mentions is

intonation, which can serve both as topicalization and a tool to keep the audience engaged. Other features are ellipses (e.g. *Back at the house...*), redundancy (repetitions serving different purposes) and others.

Some features present in the AD genre that are usually associated with written language are delayed codification (i.e. creation and presentation are spaced out), as well as delayed reception and feedback of the final audience, a high degree of planning and, in some way, permanence (the viewer can go back and re-listen to the AD however many times they want).

Given the dual nature of this special language, it remains unclear as to what the desired degree of formality and register of the text should be adopted. On the one hand, the language must be standard, “with no presence of regionalisms, localisms or idiolectal features” (Arma 2011), as well as having no metaphors and similes. This suggests that the language of AD should be neither fully colloquial nor fully literary. On the other hand, Dosch and Benecke (2004) state that in audio description “formal, written language must be avoided as this hinders a lively description following the motion and life of the movie” (2004: 24). These restrictions make it difficult to determine the appropriate register in AD.

For example, if vocabulary is considered, should Latinate verbs (e.g., *descend*, *depart*) be employed? The use of Latinate verbs has been frowned upon for at least a century, from *The King’s English* in 1906 (as quoted by Panickssery, 2023), which prescribes “Prefer the Saxon word to the Romance”, to G. Orwell’s advice, “Bad writers... are nearly always haunted by the notion that Latin or Greek words are grander than Saxon ones” in 1946 (Orwell, 1974), to P. Graham “use simple, Germanic words” (2005), and many more, suggesting that Latinate words are too complex even in written language. When it comes to content meant to be consumed through hearing, there appears to be even more reason to avoid these verbs. It can be argued, of course, that the lexical units above are not on the same level of formality as *circumnavigate* or *traverse*, for example, and are understood by most speakers with little effort. However, as Braun notes, “the sequential nature of the verbal mode seems to encourage a more complete processing of the information offered”, while “elements which are simultaneously offered by a visual image are not necessarily simultaneously processed” (2008). This indicates that AD, offering many descriptions and enumerations, while additionally compounded by music and audio cues, may require quite a lot of cognitive effort already, and any measure aimed at easing this load appears to be a good step. This

being said, whether choosing one level of formality over the other makes a difference in processing the AD by its target audience can only be established empirically.

The last section of this chapter delves into another issue associated with the need to transfer a text from a non-sequential semiotic system into a sequential one without overloading the audience.

2.8 How characters are introduced

Benecke (2014) presents the matter of character introduction as twofold. He maintains that the real problem lies in when to name the character, while all existing guidelines seem to agree on how to describe them. The German standards Benecke references give the following necessary characteristics to be named: age, hair color, height, facial expression and gestures. The AD guidelines for *Netflix* give a more extensive list: hair texture, skin color, eyes color, build, height, age description (such as late thirties, fifties, teenage, etc.), traits related to visible disabilities, etc. (*Audio Description Style Guide v2.5, 2023*). This difference is interesting, since out of the five traits cited by Benecke, only three actually describe the character, i.e. can be attributed to their permanent traits rather than what they are doing at the moment, that is, their transitory attributes (*cf. Croft's permanent and transitory states, 2012*). From this standpoint, the *Netflix* guidelines have both a wider variety of features to choose from (since it is hardly possible for an audio describer to address all of them given the time constraints) and a bigger focus on the permanent traits when describing characters. The authors raise other issues, such as that of representation and using person-first language:

...consider both the needs of the plot and the importance of representation. Description... should be done consistently for all main and relevant supporting characters that are being described, (i.e. do not single out a character because of a specific trait, describe everyone equally) and using a person-first approach (e.g. “a swimmer with one leg” instead of “a one-legged swimmer”). If unable to confirm or if not established in the plot, do not guess or assume racial, ethnic or gender identity. Instead, focus on the characters' physical attributes as described above.

(*Audio Description Style Guide v2.5, 2023*)

Indeed, the matter appears more controversial than presented by Benecke. For example, the issue of race is treated differently in AD depending on the country, but there is very little written about

it in the literature. In principle, there are four ways audio describers deal with this matter and each approach may have its own problems, as pointed out by Tair et al. (2023b):

- *Solution 1*: Not naming the race of any of the characters. This approach seems to be prevalent across AD worldwide, including *Netflix* (despite the guidelines cited above). The problem with this so-called “colorblindness” lies in the fact that, while attempting to avoid racism, it often ignores and ultimately enables it.
- *Solution 2*: Only naming the race of non-White characters, as it is done in theater AD, according to Hutchinson et al. (2020). By doing so, the audio describer deracializes Whiteness presenting it as a default, which needs no racial description, unlike other races.
- *Solution 3*: Only naming the race when it is relevant to the plot. This approach does not provide full accessibility the sighted audience has, resulting in different experiences. Plus, the importance of representation must be considered.
- *Solution 4*: Naming every character’s race. This approach may prove to be overwhelming for the audience. It also requires a lot of attention and sensitivity, given the often changing racial categories and attitudes to specific terms. Another problem with this solution was pointed out by Penny (2021), who notes that sighted viewers may not necessarily “notice” the character’s race when they are first introduced, until it becomes relevant at a later point. Therefore, making the race explicit every time may create unnecessary categorization in the viewers’ mind.

Avoidance of such “permanent” traits as race, disability, hair texture, gender (in cases where the character may be non-binary), leaves the audio describer with “transitory” traits, such as clothes, gestures and facial expressions cited by Benecke. However, this could create problems if these features have to become part of the Interim Character Fixation – the initial description used until another, more specific one, comes up in the film (Benecke, 2014). For example, in *The Life of Others* (2016), one of the characters is first described as “the man with a bald patch”, which is then changed to “the captain” (after a colleague uses this title) before being substituted by his name, Wiesler (Benecke, 2014). Relying only on transitory traits would make it more difficult to identify the character in a new scene, where they might wear different clothes and have different body language.

2.9 Specific research questions

Chapter 2 offered a comprehensive overview of the foundational concepts and principles that underpin the field of audio description. Characteristics described by AD researchers, as outlined in this chapter, serve as a base for comparing the AD tracks of two television shows. The choice of the specific categories in RQ 2 (a–e) is explained in the next chapter.

The general questions presented in Chapter 1 can be specified as follows:

RQ 1. What characteristics and observable patterns can be discerned in the AD tracks of the US “Atypical” and the Australian “Totally Completely Fine”?

RQ 2. How similar or different are the AD tracks in the two shows in terms of:

- (a) local coherence,*
- (b) global coherence (see Section 2.4),*
- (c) subjectivity and objectivity (see Section 2.6),*
- d) formality (see Section 2.7), and*
- (e) character introduction (see Section 2.8)?*

In addressing the two research questions, Chapter 3 offers a coding scheme developed for this comparative analysis.

CHAPTER 3. MATERIALS AND METHODS

3.1 Materials

The two shows selected were the US *Atypical* (*Netflix*, 2017–2021) and the Australian *Totally Completely Fine* (*Stan*, 2023), both of which are comedy dramas with each episode lasting for about 40 minutes. This pair was chosen for the comparison due to genre similarities as well as the availability of an AD track in both shows. *Atypical* was selected first due to its popularity⁸. Then, finding a match for it that would have been made in a different English-speaking country with a distinct variety of English, have an AD track and would not be on *Netflix* proved to be a challenge. Most streaming platforms do not offer audio description in the first place. Eventually, *Totally Completely Fine* (hereinafter abbreviated as *Totally*) was found on the Australian platform *Stan*⁹. More details about the two shows are presented in Table 4.

Table 4: Details about ‘*Atypical*’ and ‘*Totally Completely Fine*’

Show title	<i>Atypical</i>	<i>Totally Completely Fine</i>
Country	United States	Australia
Year	2017–2021	2023
Platform	<i>Netflix</i>	<i>Stan</i>
Genre	Comedy drama	Comedy drama
No. of seasons	4	1
Episode length	30 min	48 min
No. of episodes transcribed	7,5	4
Words in AD track per minute	61	63
Total No. of words transcribed	9,183	9,186

⁸ *Netflix original “Atypical” gains momentum and popularity.* (O’Neil, 2017). <https://concordiacourier.com/arts-entertainment/post/netflix-original-atypical-gains-momentum-and-popularity>

⁹ <https://www.stan.com.au/>

Brief plot summaries of the material that became part of the study are available in Appendix 1. It should be noted that, unlike in *Totally*, the protagonist of the US show also acts as a voiceover narrator. This seemingly would create less opportunity for AD insertions, but a comparison of 20-minute segments between the two shows revealed the word count to be almost identical (see Table 4). As a side note, the *ITC Guidance* recommend having an audio describer “of the opposite gender” when the audiovisual product already has a narrator, to avoid confusion (2000). Perhaps, it was not deemed necessary in *Atypical* due to a recognizable voice and manner of speaking of the protagonist. Both audio describers in the pair have male-sounding voices.

3.2 Data collection

In order to compare the two AD tracks they had to be transcribed first. The transcripts were done manually (by listening, pausing and typing), since there appeared to be no other way to extract the data¹⁰. This way, the corpus was produced, with each of the two shows in a separate *Google* spreadsheet. The transcripts are presented with timestamps for each scene, see Table 5 as an example.

Table 5: An excerpt of the AD transcript for ‘Atypical’

Instance ID	Episode #	Timestamp	AD track
68.	1x3 ¹¹	12:20 — 13:00	Now, Sam and Elsa in a clothing store.
69.	1x3	13:00 — 14:04	Zahid raises his arms. Zahid hugs her. She points to her head. She nods vaguely.

¹⁰ My attempts at automating this process included, for example, recording the two tracks (original and AD) from *Netflix* with a special *Chrome* extension, opening them in an audio processing app to display as two parallel tracks and manually editing out all identical segments, leaving only the ones with AD. This proved to be long and inefficient. First, the recording takes as long as the episode runs, x 2 for the two tracks. There is no other way to get the AD track since it is only available on the platform (unlike the other audio tracks, which can be downloaded by other means). Second, my initial hope was that I could find a software feature to identify and remove the identical segments automatically, but no such feature was found. Moreover, the realization that an AD track without the dialogue or context cannot be efficiently analyzed anyway rendered my efforts seemingly futile. Besides, there is no speech-to-text tool that does not require heavy editing, and organizing the data in the table with timestamps would have to be done manually in any case. Finally, the fact that the process would not get more automated with practice convinced me that transcribing the two shows completely manually was by far the easiest solution.

¹¹ Season 1, episode 3.

3.3 Annotation and comparison scheme

Transcribing the two shows was followed by a manual overview of the data and preliminary identification of potentially noteworthy features. In order to operationalize many of the factors relevant for AD that were discussed in Chapter 2, a coding scheme was constructed, see Table 6. Most features, such as formal sentence structures and vocabulary, as well as the local and global coherence, had not presented themselves as noteworthy until I started making the transcript of the second show and saw how different approaches to AD can be.

The first notes formed the foundation for constructing a coding scheme, where some of the features were grouped together under several overarching categories, while other features were considered less significant or encompassing too wide a range. In this way, the scope of the research was identified, and five categories were outlined.

Table 6 demonstrates the AD comparison scheme on which the annotation was based, with the categories and features explicated below.

Table 6: AD comparison scheme

CATEGORY	FEATURES
A. Local coherence	Referencing characters and objects from other elements of the scene (dialogue, sounds, music, voices, gestures, etc.)
B. Global coherence	Narrating flashbacks: are there specific cues used to signal the start and end of the flashback?
C. Subjectivity/Objectivity	C1 Interpretative verbs: <i>appear</i> (= seem), <i>struggle to</i> , <i>try/attempt to</i>
	C2 Named emotions
	C3 Described facial expressions
D. Formality	D1 Vocabulary: Latinate vs. Germanic
	D2 Sentence structure: conjunctions <i>as</i> and <i>before</i> ; <i>wh</i> -clauses
	D3 Sentence length
E. Character introduction	Narrating permanent and temporary character traits

The table presents five categories: Local coherence, Global coherence, Subjectivity/Objectivity, Formality and Character introduction, each including one to several features. In accordance with Section 2.4, coherence was split into two categories, A (Local coherence) and B (Global coherence). Local coherence entails continuity between the AD track and other parts of the audiovisual product, while Global coherence is utilized for an analysis of different approaches to narrating flashbacks. Category C, Subjectivity/Objectivity, comprises interpretative verbs such as *appear* in the sense of “seem”, *struggle*, as well as *try* and *attempt* (C1), named emotions (C2) and described facial expressions (C3). Next comes category D, Formality, which covers vocabulary (D1), sentence structure (D2) and sentence length (D3). Finally, Character introduction deals with narrating permanent and temporary traits of the characters when they are first introduced. The number of instances for every feature belonging to each category was quantified, and the results were further analyzed using the qualitative method.

Figure 6 shows an excerpt from the annotations done for *Totally*.

ID	Episode	Timestamp	AD track	A	*B2*	C1	*C2*	C3	D1	D2	D3	/B6/
1.	1x1	0:18 — 1:55	Now, an aerial view of surging waves crashing over rocks. /A young white woman/ sits in a bathtub, up to her chin in water. She stares "blankly" ahead. Her left hand holds onto a soap rack on a green-tiled wall. Her feet poke up above the water. Her toe nails painted black. She stares ahead. Her eyes rimmed with black eyeliner. *Images flash of a young girl.* The woman lifts one leg out of the water. Her big toe curls around the electrical cord of a hair dryer on a shelf beside the tub. *Images flash of the young girl in a car. Back to the woman.* She watches with a slight smile as she tugs the dryer cord with her toe. The dryer moves closer to the edge of the shelf, nearing the water. It teeters on the edge of the tub. The woman looks away. She closes her eyes. She lifts herself to the side of the tub and grabs her phone. The screen illuminates. The time: ten to 11 am. There's a missed call from brother 1 and a text message that reads: "Viv, please call. Emergency." A new text arrives, "I know I say this a lot, but I mean it this time." Viv lies back and answers the phone. Viv stares ahead "with a blank expression". Her eyes dart from side to side as she seems to process the news.	0	2	1	2	2	0	1	8,78	1
2.	1x1	1:55 — 2:13	Now, she steps out of the bathroom fully dressed. She walks past a bedroom, where /a white man with a black scruffy beard/ lies on his stomach. Flashes of light zap in the bathroom. She walks away.	0	0	0	0	0	0	1	9,25	1
3.	1x1	2:13 — 5:02	Viv walks down an urban side street. She passes /a burly white man with a moustache/ on the phone. He ends the call and turns to Viv. The man looks at her. His face crumples and he wraps his arms around her. He sits on a low wall and hands her a yogurt cup. She sits beside him. He hands her a little plastic spoon. She lifts the spoon of yogurt to her mouth, then looks at him. She looks at him. /A young white man/ approaches with a briefcase. Vivian looks up at him. She eats her yogurt. She man with the moustache shakes his head. He and Vivian continue to eat their yogurt. The man with the moustache begins to shake and then cry. The other man turns and walks away.	0	0	0	0	0	1	0	7,82	2
4.	1x1	5:02 — 9:09	Now, they all sit in an office, at a desk piled with papers and golfing knickknacks. A /middle-aged white man/ appears. He drives a mobility scooter. Vivian grins at him. The other man watches with a "disapproving" stare. The man with the moustache. Hendrix smiles and lowers his head. John stares "blankly". She watches as she toys with a gold chain around her neck. He hands her a piece of paper. John stares as Vivian studies the piece of paper. Hendrix nods. John bobs his head. Vivian glares at him. She turns to Hendrix. Hendrix looks at her and turns his eyes to John, who raises his eyebrows at him. Vivian stares at John. She stands up and storms off. Hendrix looks to the lawyer. Hendrix forces a smile and rushes away. John remains seated, staring ahead.	1	0	0	2	2	0	2	6,52	1

Figure 6: Annotation of the AD track of 'Totally Completely Fine'

Each of the five categories and the way they were operationalized are discussed in detail in the next section.

3.4 Procedure

3.4.1 Category A: Local coherence

This category covers the continuity between the AD and other parts of the audiovisual content, such as dialogue, audio cues, voices (when it may not be clear which character is speaking), gestures, music, etc. Due to this continuity and the way the audio describer constantly references other parts of the audiovisual product, the AD track cannot be separated from the rest of the film. A word or phrase in the AD track was coded as an instance of Local coherence if it was explicitly linked to a preceding element: image (what is depicted visually), the characters' dialogue or a sound heard in the scene. In each such instance, the word or phrase served to connect the elements in a way that made the AD track dependent on the surrounding context. Instances of Local coherence were marked as the transcripts were being made, as they immediately stood out. In such cases, the transcript of the AD track was supplemented with the element(s) it served to connect. Consider the following examples, where the AD either succeeds or fails to ensure continuity between the visual representation, the dialogue and the sound (if any).

Table 7: Examples of Local coherence

ID	Image	Dialogue	Sound	AD	Elements linked	Show
<i>a</i>	Vivian lies on the bed. Vivian sits up.	x	[sound of glass breaking]	By the sound of breaking glass, Vivian sits up.	AD + Sound	<i>Totally</i>
<i>b</i>	Dane makes a circle around his face.	Dane: And just to be clear. This...	x	His face.	Dialogue + Image (gesture)	<i>Totally</i>

		Dane: Only horror and utter confusion.				
<i>c</i>	Doug and Casey come home after a run. Doug kisses Elsa's neck. Casey pretends to vomit.	Elsa: You guys need fuel after your run? I made cinnamon buns. Doug: No, no, no. I got all the sugar I need right here.	[gag]	xxx xxx	[disturbed] Image, Dialogue, Sound	<i>Atypical</i>

Examples *a* and *b* show different ways to build continuity between the different elements: the AD and an extra (intradiegetic) sound (*a*, *Totally*), as well as the AD, the dialogue and the image, namely, a gesture, which replaces part of the dialogue (*b*, *Totally*). There are also cases where coherence is disturbed, in Braun's terms (2011). In particular, this happens in (*c*, *Atypical*): first, Casey's father, Doug, kisses her mother, Elsa, on the neck, which is not represented in the AD. Seeing this, Casey pretends to vomit to show her "disgust", which is also not relayed for the blind audience. However, the pretend retching sound is clearly heard, which leaves the AD users guessing as to who and why could have made it. The dialogue and overall context is somewhat helpful to make an assumption that Casey's parents might be kissing, and, knowing that Casey and

Sam are also in the room, perhaps, some AD users might be able to come to the right conclusion. However, the cognitive load required to use the method of elimination in a very underdescribed group scene makes the experience less enjoyable for the target audience of the AD.

3.4.2 Category B: Global coherence (narrating flashbacks and montages)

Arguing that what is shown in a film is subjective by default and the audio describer's job is to compensate for this subjectivity, Kruger suggests describing flashbacks using phrases such as "He recalls" or [something] "suddenly triggers a vivid memory", among others (Kruger, 2010). Since both shows feature flashbacks, introducing this coding index would make it possible to explore how the audio describers signal the start and the end of such instance.

3.4.3 Category C: Subjectivity/Objectivity

This section describes how subjectivity/objectivity was operationalized for the quantitative analysis. For C1 and C2 (see below), it was decided to consider the AD track more subjective if it had higher number of instances in which C1 and C2 are present. For C3, however, the higher the number of instances, the more objective was the AD, since relaying visual elements directly rather than interpreting them brings the target text (AD track) closer to the source text (audiovisual content), at least formally (see Section 2.3). Below the three features are described in more detail and examples are provided.

C1. The verb *appear* in the sense of "seem" is used in audio description as a way to interpret what is seen, for example:

(6) *She appears unsettled (Totally).*

(7) *The ocean waves appear blue in the moonlight (Totally).*

Other verbs in this subcategory are *struggle*, *try* and *attempt to*, which are utilized to present the action in a specific way: instead of saying that the character smiles, for instance, the audio describer opts for *attempts to smile*, and instead of saying that someone swallows the cereal, *struggles to swallow the cereal* is used. These words were only counted if they had this interpretative meaning.

C2. The next feature in the subjectivity/objectivity category pertains to the instances in which emotions were made explicit, i.e. named instead of described. Since guidelines explicitly prohibit inferring emotions (see Section 2.6), not many of such instances, if at all, are expected to

be found in the data. Below are two examples of this feature, while some others were listed in connection with the word *appear*.

(8) *The others stare aghast (Totally).*

(9) *John lies on his back with a pained expression (Totally).*

C3. This feature concerns facial expressions described with reference to various parts of the human face, namely, eyebrows, eyes, mouth and lips. According to guidelines, this should be the most common way to comment on emotions in AD, so both shows were expected to rely heavily on this approach. However, it was interesting to see if this feature was activated more in one of the shows. Below are examples of how the words above are used to describe facial expressions:

(10) *He chews his lip (Atypical).*

(11) *Her mouth drops open (Atypical).*

(12) *Hendrix looks at her and turns his eyes to John, who raises his eyebrows at him (Totally).*

Note that these words were only counted when they described a facial expression. If, for example, *mouth* was activated when talking about putting food into the mouth, or *eyes* were closed because the person was sleeping, these instances were not accounted for. However, if the eyes were closed in some other situations, for example “He spins slowly, his *eyes* closed and arms wide” (*Atypical*), these instances were considered, because such a facial expression could be interpreted in some other way (e.g. *He spins happily*).

3.4.4 Category D: Formality

As shown in Section 2.7, it is difficult to determine the level of formality appropriate for audio description, since this special genre is placed somewhere between spoken and written language and its specifics have not yet been described sufficiently in the literature. Formality can be assessed in various ways, one of them being the choice of specific vocabulary (coded D1). Consider these examples from the data:

(13) *She descends an outdoor staircase in front of the school and departs (Atypical).*

(14) He walks up a staircase as he takes off his paramedic shirt. John watches him go with a blank expression (Totally).

The audio describer in the first example opted for Latinate verbs in both cases, while the second AD features verbs, including a phrasal verb, of Germanic origin. This was identified as a possible feature of formality, since, according to the *Cambridge Dictionary*, “More formal vocabulary commonly involves longer words or words with origins in Latin and Greek. More informal vocabulary commonly involves shorter words, or words with origins in Anglo-Saxon” (*Formal and Informal Language*, 2024).

Besides *descend* and *depart*, other Latinate verbs of motion were selected (*ascend*, *enter*, *exit*, *approach* and *return*), as well as verbs denoting continuation of an action: *proceed* and *resume*.

The next feature in the formality category has to do with sentence structure, which is investigated by looking into the use of specific connectors: conjunctions *as* and *before* in certain environments and *wh*-pronouns introducing a subordinate clause. This feature is coded as D2 in the annotations.

Quite conveniently for AD creators, English has many ways to connect clauses that allow to cram more meaning into fewer words. However, some of these economical constructions may be considered too dense. For example, using the conjunction *as* in the way exemplified in (15) seems to be a feature of the AD language, as mentioned in Section 2.7. While, according to Bourne and Hurtado (2007), it is one of the most common sentence structures in AD, dictionaries (Merriam-Webster, Cambridge Dictionary) barely give any examples where two finite clauses with different agents are connected in this way (*A does B as A-pronoun does C* is more common). It is also said that the *as*-clause tends to be in the continuous form, which seems to not be the case in the data.

(15) The girl looks back at her book as Elsa leaves the building (Atypical).

A does B as C does D

Clauses with one agent are sometimes connected with *before*, following a non-finite phrase (gerund). According to the Cambridge Dictionary, “A non-finite clause with *before* + *ing*-form is

more formal: *Before bringing the milk to the boil, add the egg* (more formal than *Before you bring...*). Note that, unlike other more formal choices, a phrase like this is, in fact, not more economical than the conjunction *and* (16, 17). So, if it is indeed the preferred choice in AD, the most likely explanation for it would be a preference for a more formal register.

(16) *She stares at it for a moment before looking around to see if anyone is watching (Atypical).*

A does B before doing D

(17) *He ends the call and turns to Viv (Totally).*

A does B and D

The use of *wh*-clauses is also explained by the chosen register rather than time constraints, since splitting two finite clauses with different agents into two sentences would not affect the length (see 18, 19).

(18) *Doug arrives home and walks over to Elsa, who watches the kids from the kitchen.*

(Atypical).

Clause A + *wh*-clause (where, when, who, which, whose)

(19) *cf. Back to the woman. She watches with a slight smile as she tugs the dryer cord with her toe.*

(Totally).

Sentence A. Sentence B.

Lastly, sentence length (coded D3) was considered. Since each of the two annotation spreadsheets was split into scenes, this unit (scene = cell) was used for the calculations. Several steps were required to automatically calculate the average word number per sentence for the codification of D3. The formulas used are presented and explained in Appendix 2. The results of these calculations were verified via a random manual check.

The rationale behind this step was to see if either of the AD tracks had longer sentences. This was associated with the formality category because in formal language sentences tend to have one or more subordinate clauses, which may inflate their length. On the other hand, employing some formal features such as the Latinate verbs and conjunction *as* considered earlier are likely associated with a decreased word count: the Germanic equivalents for many Latinate verbs tend to

be phrasal verbs (two or more words instead of one). This metric would be useful in seeing how different choices of formality levels, among other things, might have affected the two AD tracks, given that most guidelines require AD to be “concise” (*Netflix*), “succinct” (*American Council of the Blind*) and “pared down to the essentials” (*ITC Guidance*).

3.4.5 Category E: Character introduction

The final category considered in the study has to do with naming the permanent or temporary traits when the character is first introduced. The number of character introductions was calculated for each transcript, which was followed by determining if the characteristics given for each were to be considered permanent or temporary. Age, hair color and glasses, although, strictly speaking, not “permanent” in real life, were coded as such, as well as the character’s race, while clothes were considered temporary. Gender, while considered “permanent”, was not accounted for in this category, since a character would rarely be introduced as “a person” rather than “a woman”. For this reason, instances where only the gender was given were counted separately. Since some descriptions have both permanent and temporary features, as well as gender (*/A black / woman / in a pink blouse*), the categories were named *Having permanent features*, *Only temporary features* and *Only gender*, instead of just *Permanent*, *Temporary* and *Gender*. A final *Other* category was added for the rare cases where the person is described without naming the gender.

The results of the analysis are presented in the next chapter.

CHAPTER 4. RESULTS

This chapter presents the findings of the comparative analysis conducted on the AD tracks of *Atypical* and *Totally Completely Fine*, with the five categories linked to the respective part of RQ 2: *How similar or different are the AD tracks in the two shows in terms of a) local coherence, b) global coherence, c) subjectivity/objectivity, d) formality, and e) character introduction?*

The results for RQ 1 will be presented in Chapter 5.

4.1 Local coherence

As described in Section 3.4.1, the analysis of local coherence aimed to investigate how the audio describers succeed or fail at joining all the cues available to the sighted audience in a particular scene into a unified narrative structure. In total, there were 20 instances of explicit links supporting local coherence in *Atypical*, of which 8 were considered examples of disturbed coherence. In *Totally*, this category was activated 29 times, of which one was considered a case of disturbed local coherence. A table with examples can be found in Appendix 3.

4.2 Global coherence: Narrating flashbacks and montages

This section presents different ways the audio describers approach flashbacks and montages. In line with Kruger's (2010) advice on using cues such as "He remembers", the first example from the data shows a character "recalling" the scene rather than the scene that "is shown" or "appears":

(20) *Elsa frowns as she recalls wearing the jacket and checking herself in the mirror, as well as dropping the jacket on Nick's bed (Atypical).*

In this example, Elsa (*Atypical*) first appears with a frown in the "present", and the rest of the sentence describes the flashback. This approach cannot always be applied, however, because it is not always clear which of the characters is remembering or imagining what is being shown. For instance, in the following scene the audience does not know whose mind they are in: Sam's, Julia's or both as the pair is dancing in a parking lot, Julia wearing her pajamas and a blanket (*Atypical*). In fact, the audience is not meant to know this, since, as revealed later, Sam misread the situation, too.

(21) *Julia looks up to Sam. She smiles and casts her gaze downward. Now she wears a shiny black dress as they dance in the parking lot. She leans her head on his shoulder and embraces him tightly. Wearing her blanket and sweats, Julia pulls away (Atypical).*

While it can be a challenge to introduce this imagined scene in the AD without knowing who is doing the imagining, it is interesting that the audio describer chose *now* as the cue of the scene change, as *now* may be considered insufficient or could be misinterpreted in this case.

A solution to this may be using the film jargon (e.g. “cut to...”), as recommended by McGonigle (2013; Section 2.7). The *Netflix* guidelines also urge AD writers to use “film terminology that has entered the common vocabulary... when in line with the story and/or genre” (2023). Some technical terms or language related more to how the story is organized than to the story elements was encountered in other places in the data, as shown in examples (22–25):

(22) *Now, in slow motion, the white waves of the ocean swirl together (Totally).*

(23) *The screen goes black (Atypical).*

(24) *Now, faded images: flash of sunlight through treetops (Totally).*

(25) *A distant view shows her walking onto the ledge of a very tall rocky cliff (Totally).*

The expressions “flashback to” or “in flashback” can also be considered somewhat technical, since it stands for a certain “film technique used in the creation of meaning” (McGonigle, 2013: 216). In the examples below, this is the word the audio describer opts for. The expression “back to” marks the end of the flashback:

(26) *A flashback shows Vivian as a young girl in the backseat of a moving vehicle (Totally).*

(27) *In flashback, Vivian's mother steers into an oncoming track. Back to the mediator (Totally).*

Finally, an audio describer may choose not to introduce flashbacks at all. For instance, in the following example, the highlighted part is a flashback, but its borders are not delineated:

(28) *Elsa grates a lemon peel. She and the bartender talk. Sam's dad grabs her from behind (Atypical).*

What makes such a description possible is additional assistance from the show creators so the flashbacks can be easily recognized by all audiences: extra sound effects added on the post-production stage. For instance, when a flashback appears on *Atypical*, a conventional “flashback” sound is sometimes heard. Plus, in some cases, the music changes or starts playing. A similar technique is used to signify the end of the flashback.

More examples from the two shows are available in Appendix 4, where the flashbacks and montages are highlighted in bold. Overall, *Atypical* features 10 flashbacks/montages, of which 2 are explicitly introduced by the audio describer, whereas *Totally* has 19 flashbacks/montages, all of which are introduced by the audio describer.

4.3 Subjectivity/Objectivity

4.3.1 Interpretative verbs (*appear, struggle, try, attempt*)

Table 8 shows the number of times the selected interpretative verbs were found in each show. As can be seen, there are more interpretative verbs in *Atypical* and some reasons for this difference are discussed in Chapter 5. Examples (29–32) from the data show how each of the interpretative verbs serves to construe the message in a particular way.

Table 8: *The activation of the interpretative verbs in the data*

	Instances in <i>Atypical</i> (US)	Instances in <i>Totally</i> (Australia)
<i>appear</i>	0	7
<i>struggle</i>	0	9
<i>try</i>	2	2
<i>attempt</i>	2	0
Total	4	18

(29) She appears to collect her thoughts (*Totally*).

(30) John struggles to answer the phone (*Totally*).

(31) He tries to smile again (*Atypical*).

(32) She bites her lip and then attempts to smile (*Atypical*).

4.3.2 Named emotions

In this subsection, both ascribing an emotion and using words such as *blankly* (which suggests a lack of expression) by the audio describer are considered a subjectivity/objectivity feature. The findings in this feature are as follows: *Atypical* has 4 named emotions (all of them with *blankly*), while the AD in *Totally* mentions 15 emotions, including *blankly*, *aghast*, *disapproving*, *pained (expression)*, *unsettled*, *distressed* and more.

4.3.3 Described facial expressions

Table 9 shows the number of times the selected words referring to parts of the human face were mentioned with the purpose to describe a facial expression in the two shows. As can be noted, *eyes* were referenced by the two audio describers nearly equally, but there are no similarities beyond that. Starting from the *eyebrows*, it is obvious that the US AD describes facial expressions differently from the Australian one, while the contrast between the activation of *lips* and *mouth* between the shows is also quite illuminating.

Table 9: The words referring to parts of the human face in the data

	Instances in <i>Atypical</i> (US)	Instances in <i>Totally</i> (Australia)
<i>eye/eyes</i>	37	36
<i>brow/brows/eyebrow/eyebrows</i>	29	8
<i>lip/lips</i>	27	0
<i>mouth</i>	23	0
Total	116	44

4.4 Formality

4.4.1 The use of Latinate vs. Germanic verbs in the AD

The findings for this feature are displayed in Table 10. Notably, the US *Atypical* exhibits more Latinate verbs than its Australian counterpart (71 against 15). In examples (33–39) from the corpus, these Latinate verbs are contrasted with their Anglo-Saxon equivalents from the dataset.

Table 10: The Latinate verbs in the data

Latinate verbs	Instances in <i>Atypical</i> (US)	Instances in <i>Totally</i> (Australia)
Descend	6	0
Ascend	1	0
Depart	4	0
Enter	18	3
Exit	3	0
Return	11	4
Approach	17	8
Resume	7	0
Proceed	4	0
Total	71	15

(33) a. Sam heads towards the porch and ascends the front steps (*Atypical*).

b. Anna watches as he goes up the stairs (*Totally*).

(34) a. His dad enters and smiles (*Atypical*).

b. Later, Vivian walks into the dim-lit living room (*Totally*).

(35) a. Elsa turns and exits the room, as Casey hugs herself (*Atypical*).

- b. Then she finishes a glass of wine and walks out of the room (Totally).
- (36) a. The bartender approaches Elsa (Atypical).
b. Vivian steps closer to him until they're face to face (Totally).
- (37) a. Elsa returns to the bar to retrieve her credit card (Atypical).
b. Vivian walks back into the house (Totally).
- (38) a. Sam proceeds to climb through the window as Doug gets out of the car (Atypical).
b. Alejandro takes his hand and goes to kiss it (Totally).
- (39) a. He turns to leave and she resumes locking up her bike (Atypical).
b. John turns and faces her. He goes back to scraping dishes (Totally).

4.4.2 Phrases with *as* and *before*, *wh*-clauses

Table 11 presents the number of times the three constructions (40) are activated in the pair under study.

(40) A does B as C does D

A does B before doing D

Clause A + wh-clause (*where, when, who, which, whose*)

Table 11: Phrases with 'as' and 'before', *wh*-clauses

	Instances in <i>Atypical</i> (US)	Instances in <i>Totally</i> (Australia)
as	109	50
before	43	2
wh-closes	18	8
Total	170	60

The findings reveal another discrepancy between the two AD tracks, demonstrating the US show to be more reliant on these specific constructions.

4.4.3 Average sentence length

The average length of an AD sentence was calculated in *Google Sheets* with the help of three formulas, as described in Section 3.6.1 and Appendix 2. The results revealed that the average

sentence length in *Atypical* is 7.9 words, and in *Totally* it comes to 7.8 words per sentence. Therefore, despite the different approaches to formality, sentence structure, complexity and other features addressed below, the AD tracks in the two shows exhibited similar pacing.

4.5 Character introduction

As observed in 2.8, when characters in an AD track are introduced, the audio describer may choose to present either their temporary or permanent features. The number of characters introduced in the two shows turned out to be quite similar: 24 for *Atypical* and 22 for *Totally*. Table 12 shows some examples of how permanent and temporary characteristics are used in the dataset.

Table 12: Permanent and temporary features in character description (examples)

Character	Having permanent features	Only temporary features
Sam (<i>Atypical</i>)	A dark-haired teen	
Sam's date (<i>Atypical</i>)	A red-headed girl who wears glasses	
Bailey (<i>Atypical</i>)		A girl in white
Bailey's boyfriend Henry (<i>Atypical</i>)		A boy in a leather jacket
Coach Briggs (<i>Atypical</i>)		A man in a grey polo
Vivian (<i>Totally</i>)	A young White woman	
Hendrix (<i>Totally</i>)	A burly White man with a mustache	
Richie (<i>Totally</i>)	A man with a black scruffy beard	
Lawyer (<i>Totally</i>)	A middle-aged White man	
Amy (<i>Totally</i>)		A woman in a wedding dress

Table 13 shows that the audio describer in *Atypical* opts for temporary as well as permanent characteristics, while *Totally* describes most characters by giving their permanent features, such as

race (see examples in Table 12). Out of the 24 characters in *Atypical*, 8 were described using permanent traits and 9 only through their temporary ones. Instances where only the character's gender was given accounted for 7 times. In *Totally*, where 22 characters are introduced, the inclusion of permanent features accounted for 17 cases (since, unlike *Atypical*, most characters' race was mentioned). There was one case with only the gender given and two instances where the characters were referred to as *a mourner* and *the priest* ("Other" in Table 13).

Table 13: Permanent and temporary features in character description (Results)

Show	Having permanent features	Only temporary features	Only gender	Other	Total
<i>Atypical</i>	8	9	7	0	24
<i>Totally</i>	17	2	1	2	22

CHAPTER 5. DISCUSSION

This chapter addresses both research questions by interpreting the findings presented in Chapter 4, generalizing as much as possible from the specific versions of the questions formulated at the end of Chapter 2, to the more general initial versions from Chapter 1. Sections 5.1–5.2, address RQ 1, while Sections 5.3–5.7 sum up the differences and similarities between the US and the Australian shows in terms of their AD tracks, as asked by RQ 2.

5.1 The main features of AD

The first research question asks what the main features and current trends of AD are. As established in Chapter 2 and confirmed by the corpus analysis, some of the prominent features to be addressed when looking into an AD track are local coherence, global coherence (in particular, flashbacks and montages), subjectivity/objectivity (in particular, interpretative verbs, named emotions and described facial expressions), formality (Latinate verbs and other vocabulary choices which can be investigated further; grammatical structures, which are complex or uncommon in the general language; average sentence length), and character introduction (mentioning permanent or temporary features of the characters).

5.2 Current trends in AD

While analyzing the material, a number of trends were revealed, some of which are outlined below.

5.2.1 Some facial expressions are underdescribed

In some cases a particular facial expression can be underdescribed, since it lacks distinctive visual cues such as “widened eyes” or “raised eyebrows”. For example, when the therapist Julia (*Atypical*) is reading her boyfriend’s breakup note, she is visibly on the brink of crying, but her face is not described in the AD, see Figure 7. Similarly, when the protagonist’s sister Casey (*Atypical*) sees her mother kissing a man that is not her father, this comes completely unexpected for the girl. Her face expresses shock and disbelief, but instead, the AD simply says, “Casey watches” (see Figure 8).



Figure 7: “She moves toward it and finds a note that reads...” *Atypical*; 1x5; 24:2–24:49



Figure 8: “Casey watches”. *Atypical*; 1x7; 23:51–26:05

As for emotions that are left out entirely, consider Figure 9, where the character Alejandro (*Totally*) says “That’s perfect!”, while his turned face displays a comical grimace, signaling that “the beige corduroy pants” are anything but perfect in his opinion. However, the AD appears to only have time to describe the item of clothing in question and not Alejandro’s face.

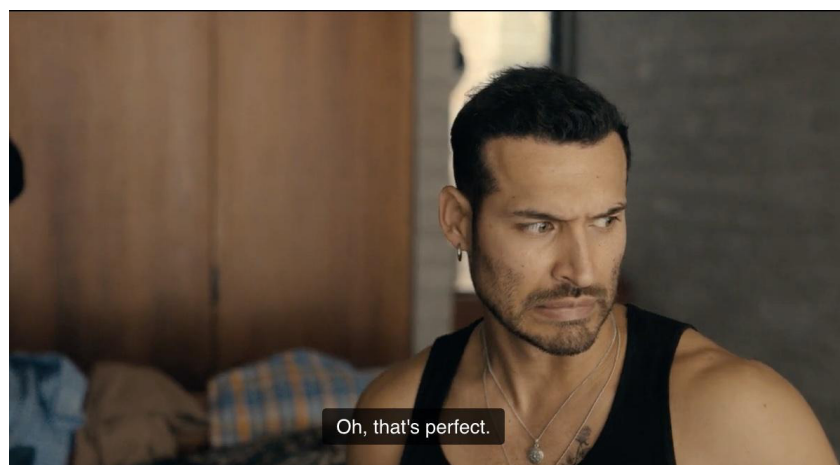


Figure 9. Dialogue: “That’s perfect!” AD: None. Totally Completely Fine; 1x2; 26:22–28:05

And lastly, in one of the scenes featuring Casey with her mother Elsa (*Atypical*), the AD says, “Elsa smiles and Casey smiles back”, but the smile on Casey’s face is visibly sarcastic, which remains inaccessible to the blind audience (Figure 10).



Figure 10: “Casey smiles and Elsa smiles back”. Atypical; 1x6; 8:50–10:09

The choice to underdescribe or entirely omit a facial expression may be associated with the audio describers’ efforts to follow the guidelines prohibiting any subjective interpretation (in particular, US guidelines are known for this, see Section 2.2). Subtle movements, such as a small twitch of the eyebrow or a faint change in the eyes, are not always easy to describe in a way that is clear and objective. When audio describers feel an expression would require subjective interpretation to be

meaningful, they might prioritize their interpretation-free approach over accuracy. This could indicate that the streaming platform guidelines and the way they are enforced on the AD team might be overly restrictive. It also suggests lack of proper communication between the contractor and the client, otherwise such cases would have been negotiated as exceptions from the strict objectivity rule. Instead, the way these descriptions are done seems to be a one-size-fits-all approach without considerations for specific circumstances.

5.2.2 Using “frown” as a general term for various facial expressions

Sometimes the characters in the dataset are described as “frowning” when they are showing facial expressions more complex than that. For example, the protagonist of *Atypical*, Sam “frowns” after his girlfriend Paige breaks up with him (Figure 11), and his sister Casey “frowns” when her best friend betrays and leaves her (Figure 12). Figures 13 and 14 show two more cases with their respective descriptions. These four instances indicate that the word *frown* a) is not always used in accordance with the dictionary definition¹² and b) is applied to facial expressions that bear no similarities to each other. This approach, Bittner warns, may present different emotions and situations as “exchangeable”, since the audio describers always default to the same descriptions (Bittner, 2014).



Figure 11: “Sam sits cross-legged, looking around and frowning.” *Atypical*; 1x5; 20:41–21:16

¹² “to bring your eyebrows together so that there are lines on your face above your eyes”, Cambridge Dictionary)



Figure 12: "She frowns." Atypical; 1x6; 14:21–15:29



Figure 13: "Sam frowns and breaks eye contact." Atypical, 1x2; 02:30–3:18



Figure 14: "Elsa frowns as she recalls wearing the jacket..." Atypical, 1x4; 18:15–21:34

5.2.3 Using repetitions in AD

Sometimes, rather than opting for a synonym or a new turn of phrase, audio describers may utilize repetition. In certain cases, such a choice seems rather deliberate. Three purposes of repetition in AD may be outlined.

First, naming recurring objects consistently to build immediate recognition. Audio describers may avoid synonyms when announcing surrounding objects, such as *the beach house*, *the coffee table*, *the trellis archway*, *the dense vegetation (Totally)*, etc. This is one of the strategies, which should help reduce the cognitive load for the recipients. The *Netflix* guidelines, for instance, state: “visual elements (e.g., the naming of locations) should remain consistent within the description for the entirety of the content and across episodes/seasons. A glossary should be created listing common descriptors.”

Second, reflecting repetition in the visual content, which can be approached in different ways. Such repetitions will not necessarily be reproduced word for word, since the visual elements themselves are often not identical. In cases like this, words such as *again*, *as well*, *repeat* etc. are used, along with verbs with the *re-* prefix. In this event, the wording may be either preserved or changed.

(41) *She closes the door, as Casey rubs her head. Again, she comes back. She closes the door. She reopens it again. (Atypical)*

(42) *Elsa shuts the laptop quickly, gets up and grabs a pair of shoes to throw in the dryer. Then sits next to it as it runs... Once again, Elsa grabs the sneakers and throws them into the dryer. (Atypical)*

(43) *Sam reenters the house. Sam retakes his seat. (Atypical)*

(44) *He turns to leave and she resumes locking up her bike. Sam turns back. Again, he turns away, but again, he returns. (Atypical)*

The audio describer can even make a comparison between similar scenes, for example:

(45) *Then, Casey runs down the same asphalt path that she and Doug ran on earlier. This time, on her own. (Atypical)*

Third, emphasizing prominent character features, which helps the audience piece together various traits of the characters as they become revealed, accumulated and transformed (Bal, 1997; Vercauteren, 2016). For example, every time John’s house is shown on *Totally*, it is described as *modern*. This is one of John’s prominent characteristics, since he appears to use external signs of success in an attempt to cover up his insecurities, see (46–49):

(46) *In a modern kitchen, John... (Totally)*

(47) *A towering plant grows in the lobby of John and Alejandro's spacious modern home. (Totally).*

(48) *He disappears up a modern staircase with black railings (Totally).*

(49) *Meanwhile, at John and Alejandro's modern home... (Totally).*

More examples of how repetition is used in AD to reflect repetition in the visual content and to narrate prominent character traits are available in Appendix 5. Since such uses of repetition help establish links between different scenes, it may be considered part of global coherence and become an object of study in its own right in future research.

5.3 Local coherence

Section 4.1 showed that *Totally* both has more explicit links within a scene than *Atypical* and fewer cases where the local coherence is disturbed, e.g., when an important item in the frame or the source of a sound is not named. These results present the AD track in *Totally* as more beneficial for the users as it ensures clearer and more consistent storytelling. Consequently, this enhances the overall viewing experience for visually impaired audiences by providing them with a more comprehensible and engaging narrative.

5.4 Global coherence

According to the findings in Section 4.2, *Atypical* features fewer flashbacks and montages and most of them are described without specific cues signaling their start and end. By contrast, the audio describer in *Totally*, which has almost twice as many flashbacks, introduces all of them, whether by saying the word “flashback” or otherwise. This is another aspect showing the AD track in *Totally* as seemingly more accommodating to the needs of visually impaired audiences. It should

be noted that lack of explicit verbal introduction does not mean the blind audience does not receive any indication of a flashback whatsoever, since there could be additional sound effects making it clear that a shift has started or ended. However, these sound cues may still be ambiguous and whether blind audiences would be able to correctly identify a flashback (as well as its end) in 100% of the cases should be tested empirically. Although some sounds could be linked to a flashback (e.g. there are “flashback sounds” in sound banks), they are not used in all cases, and other factors may get in the way of AD users’ comprehension and enjoyment.

5.5 Subjectivity/Objectivity

Section 4.3 showed that *Totally* had more (18 vs. 4) interpretative verbs (namely, *appear*, *struggle*, *try* and *attempt*). Such a difference may be associated with the AD track in the Australian show being more inclined towards subjectivity. It can also be the case that *Totally* featured more unsuccessful actions (“struggling”) for a comedic effect. Although the shows do share a genre, there are distinct elements comprising a comedy in different countries, in which case the AD in *Totally* may simply be reflective of the visual part of the product.

Totally has also turned out to have more named emotions and fewer described facial expressions. This serves to highlight a more subjective approach by the audio describer of the Australian show when compared to its US counterpart. One of the reasons for this could be that the Australian AD standards might have been made closer to the British ones, which, as discussed in Section 2.2, allow more interpretation. Another possible explanation is that there are no firmly established AD practices in Australia, so, perhaps, no strict rules regarding the naming of emotions and description of facial expressions have been formulated yet. However, as observed in Section 2.2, China, while also being new to AD, very early developed a policy that prescribes objectivity in all cases. It should also be noted that the two AD tracks may not be representative of the audio description practices in the United States and Australia in general. Perhaps, further research will shed more light on the reasons for these differences and allow to make generalizations about the AD in the two countries.

5.6 Formality

The results described in Section 4.4 showed that *Atypical* has more formal features than *Totally*. Namely, it exhibits more Latinate verbs (71 vs. 15) and complex grammatical structures (170 vs.

60). However, the two shows turned out to be quite similar when it came to the average sentence length, with *Atypical* having 7.9 words and *Totally* 7.8 words per sentence.

It is difficult to say whether these dissimilarities stem from differences between the varieties of English or happen due to different approaches to AD. However, these particular choices (Linate verbs and the three constructions) have not been found in sources discussing the distinctive features of the US and the Australian varieties (e.g., Trudgill & Hannah, 2013). Wardhaugh and Fuller observe:

Today, Standard English is codified to the extent that the grammar and vocabulary of English are much the same everywhere in the world: variation among local standards is really quite minor, being differences of ‘flavor’ rather than of ‘substance,’ so that the Singapore, South African, and Irish varieties are really very little different from one another so far as grammar and vocabulary are concerned.

(Wardhaugh & Fuller, 1986 :35)

Given that Standard English is the “variety of English which is usually used in print, and which is normally taught in schools and to non-native speakers learning the language” (Trudgill, 1995:5), it stands to reason that it must incorporate formal language. This means that the variances found in the two AD tracks are most likely associated with the different approaches to making AD — whether between the two shows or between the two countries is to be established by further research.

5.7 Character introduction

As shown in Section 4.5, there is a difference in the ways characters are described in *Atypical* and *Totally*. While in *Atypical* the distribution among the categories “Having permanent features”, “Only temporary features” and “Only gender” is almost equal, *Totally* exhibits a clear preference for the first category. It is mostly associated with the fact that the Australian audio describer names the race or ethnicity of most characters, while their US counterpart does not do that for any of the characters, relying on the description of their clothes instead.

CHAPTER 6. CONCLUSIONS

Audio Description is developing rapidly as a practice, but research into how it is actually carried out is lagging. To contribute to closing this gap, this thesis compared two AD tracks from different varieties of English. The aim of the study was twofold: to identify features and trends in AD as it is practiced (as opposed to comparing guidelines) and pinpoint differences and similarities between a US AD track for the show *Atypical* and an Australian one for the show *Totally*.

A preliminary analysis of the AD transcripts for these shows revealed a number of features which were eventually operationalized into a comparison scheme. Some of these features had been studied by researchers focusing on AD, while others had to be approached by adapting findings from adjacent areas. As a result, the most prominent and consistent features formed five categories: Local coherence, Global coherence, Subjectivity/Objectivity, Formality and Character introduction. By making a careful comparison between the two transcripts using this scheme, both more systematic and less systematic observations could be made.

The less systematic observations were considered “trends” in AD. For instance, a number of facial expressions represented visually appeared to be underdescribed or completely left out by the audio describers. Examples show that it is not always a matter of time constraints, but, possibly, a choice associated with the complex nature of some expressions. Perhaps, the audio describer did not want to address the Subjectivity/Objectivity category by either naming the emotion, or else describing the facial expression, given that no visible movement of the eyes, eyebrows or other parts of the character’s face could be detected. Another related trend has to do with using the word *frown* for such subtle emotions, thereby extending its dictionary definition to cover all ambiguous cases. The third trend, repetition, was identified, with three of its purposes: naming recurring objects consistently to build immediate recognition, reflecting repetition in the visual content and emphasizing prominent character traits.

The more systematic observations concerning the five categories can be regarded as “main features” in the sense of RQ 1 and utilized in RQ 2 for the comparison between the two AD tracks. An analysis of local coherence revealed that the Australian show had more explicit links within a scene than its US counterpart, while also featuring fewer cases where the local coherence was disturbed. Examining global coherence in the two AD texts showed that the Australian *Totally* contained more flashbacks and the borders of each were made explicit for the viewer. By contrast,

the visually impaired audience of the US *Atypical* sometimes had to use secondary cues to identify the flashbacks, such as a different music tempo or the context.

Comparing the two AD tracks in terms of Subjectivity/Objectivity revealed that the Australian show had more interpretative verbs (*appear, struggle, try* and *attempt*). While this could be a sign of a more subjective approach, it could also be the case that *Totally* depicted its characters' as awkward more often than *Atypical* did, which was reflected in the AD track. The Australian show also had more named emotions when compared with the US *Atypical*, suggesting that the Australian audio describer might have had more freedom while interpreting the characters' expressions. At the same time, when it came to describing the faces by referring to the characters' eyes, eyebrows, mouths and lips, only the word *eyes* was activated with a similar frequency in the two AD tracks, while *eyebrows* were mostly referred to in *Atypical*. What is more, *mouth* and *lips* were found exclusively in the US descriptions, which could either mean that emotions were narrated in more detail there or that such descriptions were used in places where the Australian AD opted for *naming* the emotion.

More focused research is called for to determine the reasons for these differences. For example, could it be that such an approach was also influenced by the content of the two shows? Given that *Atypical* centers around a boy with autism, perhaps, the AD team wanted to be extra objective and careful so as to not come across as “evaluative” of the protagonist’s “atypical” behaviors? On the other hand, presenting Sam in a more interpretative way might have fostered more empathy towards him from the audience. For example, Hecht (2022) suggests that utilizing metaphors (i.e., “subjective language”, per Luque & Gallego, 2019) commonly employed by individuals with autism could provide deeper insights into their emotions and experiences. Whatever the case might be, it seems that the approach to AD in the US show aims to avoid any descriptions which could be considered “sensitive”, whether it is a health condition or an ethnicity, as can be concluded from the findings in the character introduction category as well.

An analysis of formality showed that the US audio describer used more Latinate verbs, while their Australian colleague opted for Germanic phrasal verbs instead. Besides, the two shows differ in the number of complex grammatical structures, which had been selected as part of this study (see Section 4.4.2). Each of the constructions was activated more in the US show, which, once again, presents this AD track as the more formal between the two. A reason for two AD tracks to differ in the level of formality could be, for example, different tones and styles of the shows

themselves. However, in this case both belong to the same genre and most characters in *Atypical* use the same informal language as the characters in *Totally* do. Unlike the Australian AD, the AD in *Atypical* appears to be in contrast with the level of formality in most dialogues. This difference in tone might have been introduced intentionally, so that the audio describer's narration would stand out. As noted in 3.1, *Atypical* has a voiceover narrator, while *Totally* does not. This means that the visually impaired audience of the US show has to distinguish among three narrations delivered by male voices. By contrast, the protagonist of *Totally* is a woman and there is no voiceover. Besides, the fact that Sam has autism could also be taken into account in this regard. As noted by Vogindroukas et al. (2022), "ASD individuals' verbal discourse also often includes formal words", which is an interesting point to consider. Indeed, while most characters in *Atypical* speak in an informal language, the audio describer might have chosen to present Sam's perspective using vocabulary which he himself would use. This perspective contradicts the one about the audio describer "standing out", since, in this case, all three narrators would have the same tone. Given that analyzing the formality of the dialogue was beyond the scope of this research, these are just some considerations that could potentially be taken into account when explaining the differences in the vocabulary and grammar presented here. Notably, the third feature related to the category of formality, the average sentence length, revealed a very small difference of 0.1 words per sentence. This is the only feature in the study that presented the two AD tracks as similar.

The final category, Character introduction, was used to investigate which traits the two AD tracks referred to when describing the characters. The analysis revealed that the Australian audio describer relied on "permanent" features more often than the US one, whose description exhibits a more equal distribution across the four subcategories (*Having permanent features*, *Only temporary features*, *Only gender* and *Other*). This difference can be explained by the fact that the AD in *Totally* names the race of each character and mentions their approximate age more often, while *Atypical* never comments on the race and only indicates the characters' age by referring to some male ones as "a teen" and their female peers as "a girl". It should be noted that the findings for *Atypical* contradict the *Netflix* guidelines regarding character description (see Section 2.8).

Perhaps, this section in the guidelines was updated after the AD track was made¹³. As suggested above, some “permanent” features could be considered sensitive and therefore avoided by the audio describer of *Atypical*. The fact that the cast of the show is quite diverse but the visually impaired audience is unable to appreciate it exposes an independent rather than collaborative nature of the relationship between the audio describer as a contractor and *Netflix* as a client (see Section 2.2). This contradiction is not observed in *Totally*, where the show creators and the audio describer display the same approach to diversity.

While the differences and similarities of the two television shows under study may not reflect the state of AD in the two respective countries, they can still shed some light on accepted practices. The coding scheme proposed in the study can facilitate future research, where the identified categories may be developed further. For instance, other shows on *Netflix* may be compared to see if the guidelines are followed and more US shows may be examined to see if the findings for *Atypical* hold for audio described content on platforms other than *Netflix*. Similarly, more Australian shows may be looked into to see if different platforms have other approaches to the categories identified here. As noted in 5.4, empirical research is needed to find out the best approaches to narrating flashbacks. Similarly, studies with participants are called for to reveal how explicit the links between local coherence elements should be so that visually impaired people could get the most out of the AD track. On top of that, further research is required to understand whether increased formality in AD is associated with increased cognitive load.

In sum, the thesis has investigated some of the main features and trends in current AD practices in the English language and compared the AD tracks of two television shows in different English varieties. This was done with the help of a coding scheme developed as part of this study. The scheme with its categories and features can be applied and adapted in future research, contributing to the advancement of comparative AD studies.

¹³ The last update of the guidelines was on 27 April 2023. It is unknown which sections were updated and when the AD track was added to the first season of *Atypical*.

REFERENCES

- Alattar, H. S. (2021). Arabic Audio Description: Challenges and possibilities. <https://www.ata-divisions.org/AVD/arabic-audio-description-challenges-and-possibilities/>
- Arma, S. (2011). The language of filmic audio description: A corpus-based analysis of adjectives (Doctoral dissertation). Università degli Studi di Napoli Federico II. <http://www.fedoa.unina.it/id/eprint/8740> doi:10.6092/UNINA/FEDOA/8740
- Arma, S. (2012). "Why can't you wear black shoes like the other mothers?" Preliminary investigation on the Italian language of audio description. In E. Perego (Ed.), *Emerging topics in translation: Audio description* (pp. 37–55). Trieste: EUT.
- Audio Description Coalition* (2009). Audio Description Standards. Retrieved from https://www.perkins.org/wp-content/uploads/elearning-media/adc_standards.pdf
- Audio Description in Australia – Words Becoming Pictures*. (n.d.-b). Retrieved from <https://audiodescriptionau.com.au/>
- Audio Description Style Guide v2.5*. (2023). Netflix: Partner Help Center. <https://partnerhelp.netflixstudios.com/hc/en-us/articles/215510667-Audio-Description-Style-Guide-v2-5>
- Bal, M. (1997). *Narratology: An introduction to the theory of narrative*. 2nd ed. Toronto, University of Toronto Press.
- Baldry, A. & Thibault, P. J. (2005), *Multimodal transcription and text analysis: A multimedia toolkit and coursebook*. London: Equinox Publishing Ltd, 1–56.
- Benecke, B. (2014). Character fixation and character description: The naming and describing of characters in *Inglourious Basterds*. In A. Maszerowska, A. Matamala, & P. Orero (Eds.), *Audio Description: New perspectives illustrated* (pp. 141–157). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Bellringer, H. (2015, September 25). James Bond: the greatest example of music as marketing. *The Drum*. <https://www.thedrum.com/opinion/2015/09/25/james-bond-greatest-example-music-marketing>

- Bittner, H. (2014). Audio description guidelines – a comparison. Retrieved from https://www.uni-hildesheim.de/media/_migrated/content_uploads/AD_Guidelines_Comparison_-_Read.pdf
- Bloodaxe Books. (2021, March 23). *Launch by Maria Stepanova & Pia Tafdrup with Sasha Dugdale* [Video]. YouTube. <https://www.youtube.com/live/qcCsXmO15Ac?t=2949s>
- Bourne, J., & Hurtado, C. J. (2007). From the visual to the verbal in two languages: a contrastive analysis of the audio description of *The Hours* in English and Spanish. In *BRILL eBooks* (pp. 175–187). https://doi.org/10.1163/9789401209564_013
- Braun, S. (2008). Audiodescription research: State of the art and beyond. *Translation Studies in the Third Millennium*, 6, in press.
- Braun, S. (2011). Creating coherence in audio description. *Meta*, 56(3), 645–662. <https://doi.org/10.7202/1008338ar>
- Braun, S. (2021). Audio Description from a discourse perspective: a socially relevant framework for research and training. *Linguistica Antverpiensia New Series – Themes in Translation Studies*, 6. <https://doi.org/10.52034/lanstts.v6i.197>
- Brown, G., Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press.
- Chemaly, F. (2021). A picture for the blind in a visual world. *The Independent on Saturday*. <https://www.iol.co.za/ios/news/a-picture-for-the-blind-in-a-visual-world-30d7d21c-6d51-46da-b41e-cbdf22cfa3ea>
- Connell, B. R., Jones, M., Mace, R., Mullick, A., Ostroff, E., Sanford, J., Steinfeld, E., Story, M., & Vanderheiden, G. G. (1997, April 1). The principles of universal design. Retrieved from https://projects.ncsu.edu/design/cud/about_ud/udprinciplestext.htm
- Croft, W. A. (2012). *Verbs: aspect and causal structure*. <http://ci.nii.ac.jp/ncid/BB09602240>
- Cussen, E. (2013b, January 14). New captioned cinema technology launched in Italy. *Media Access Australia*. https://www.mediaaccess.org.au/latest_news/general/new-captioned-cinema-technology-launched-in-italy
- Diget, I.S.K. (2019). *Intersemiotic translation from film to audio description: A cognitive semiotic approach*. (MA Thesis) Lund University.
- Dosch, E., & Benecke, B. (2004). *Wenn aus Bildern Worte werden : durch Audio-Description zum Hörfilm*. <https://lib.ugent.be/en/catalog/rug01:002226337>

- Formal and informal language. (2024). In *Cambridge Grammar*.
<https://dictionary.cambridge.org/grammar/british-grammar/formal-and-informal-language>
- Franca, S. (2016). Watching films through words: A comparison of the Italian and English audio descriptions of *The Silence of the Lambs* [MA thesis, Università degli Studi di Padova].
<http://tesi.cab.unipd.it/52789/>
- Fryer, L. (2018). The independent audio describer is dead: Long live audio description! *Journal of Audiovisual Translation*, 1(1), 170–186. <https://doi.org/10.47476/jat.v1i1.52>
- Gaballo, V. (2012). [Exploring the Boundaries of Transcreation in Specialized Translation](#) (PDF). ESP Across Cultures.
- Gibson, G. (2021). Words worth 1,000 pictures. *Chicago Tribune*. Retrieved from
<https://www.chicagotribune.com/news/ct-xpm-1989-12-29-8903210625-story.html>
- Graham, P. (2005). *Writing, briefly*. Retrieved from: <https://paulgraham.com/writing44.html>
- Halliday, M. a. K. (1978). *Language as social semiotic : the social interpretation of language and meaning*. <https://ci.nii.ac.jp/ncid/BA0365055X>
- Hecht, J. (2022, January 7). *How to write an autistic character*. Retrieved on February 12, 2024, from <https://allwritealright.com/how-to-write-an-autistic-character/>
- Hutchinson, R., Thompson, H., & Cock, M. (2020). Describing diversity: An exploration of the description of human characteristics and appearance within the practice of theatre audio description. <https://vocal-eyes.co.uk/describing-diversity-report-published/>
- ITC (Independent Television Committee) (2000): ITC guidance on standards for audio description. Retrieved from <http://msradio.huji.ac.il/narration.doc>
- Kruger, J. (2010). Audio narration: re-narrativising film. *Perspectives: Studies in Translatology*, 18(3), 231–249. <https://doi.org/10.1080/0907676x.2010.485686>
- Langacker, R. W. (2015). Construal. In E. Dąbrowska and D. Divjak (Ed.), *Handbook of Cognitive Linguistics* (pp. 120–143). De Gruyter eBooks. <https://doi.org/10.1515/9783110292022>
- Luque, M. O., & Gallego, S. S. (2019). Paintings to my ears: A method of studying subjectivity in audio description for art museums. *Linguistica Antverpiensia New Series - Themes in Translation Studies*, 17. <https://doi.org/10.52034/lanstts.v17i0.472>
- Mazur, I., & Chmiel, A. (2012). Audio description made to measure: Reflections on interpretation in AD based on the Pear Tree Project data. In *BRILL eBooks* (pp. 173–188). https://doi.org/10.1163/9789401207812_011

- McGonigle, F. (2013). Audio description and semiotics: The translation of films for visually-impaired audiences. https://epubs.surrey.ac.uk/805168/2/Thesis_Final_copyright%20material%20removed.pdf
- Nida, E. A. (1964). *Toward a science of translating*. <https://doi.org/10.1163/9789004495746>
- Nida, E. A., & Taber, C. R. (1982). *The theory and practice of translation*. <https://doi.org/10.1163/9789004669147>
- Ofcom (2006). Code on television access services. Retrieved from <http://stakeholders.ofcom.org.uk/binaries/broadcast/other-codes/ctas.pdf>
- Orwell, G. (1974). Politics and the English language. In *Palgrave Macmillan UK eBooks* (pp. 423–437). https://doi.org/10.1007/978-1-349-23769-2_19
- Panickssery, A. (2023, March 30). Germanic vs Latinate Words in Writing. *Arjun Panickssery*. <https://arjunpanickssery.substack.com/p/germanic-vs-latin-words-in-writing>
- Penny, J. (2021, December 16). Audio-describing diversity at ITV. *Medium*. <https://jonathan-penny.medium.com/audio-describing-diversity-or-addressing-the-elephant-in-the-room-5d731c0ea8d3>
- Piety, P. J. (2004). The language system of audio description: An investigation as a discursive process. *Journal of Visual Impairment & Blindness*, 98(8), 453–469. <https://doi.org/10.1177/0145482x0409800802>
- Pujol, J. (2007). Audio description or audio narration? That is the question. Paper presented at the MuTra Conference 2007, Vienna. Retrieved from http://www.euroconferences.info/2007_abstracts.php
- Redden, A. (2015b, April 21). Swedish app massively expands audio described cinema. *Media Access Australia*. https://mediaaccess.org.au/latest_news/web/swedish-app-massively-expands-audio-described-cinema
- Pujol, J. & Orero, P. (2007). Audio Description Precursors: Ekphrasis and Narrators. In *Translation Watch Quarterly* 3(2): 49–60.
- Remael, A. (2005). Audio description for recorded TV, cinema and DVD. An experimental stylesheet for teaching purposes. Retrieved from: <https://bibbase.org/network/publication/remael-audiodescriptionforrecordedtvcinemaandddvdxperimentalstylesheetforteachingpurposes-2005>

- Remael, A., Reviers, N. & Vercauteren, G. (2015). Pictures painted in words: ADLAB Audio Description guidelines. Trieste: EUT.
- RNIB & VocalEyes (2003). The talking images guide. Museums, galleries and heritage sites: Improving access for blind and partially sighted people. *Royal National Institute for the Blind*.
- Rohrbach, A., Torabi, A., Rohrbach, M., Tandon, N., Pal, C., Larochelle, H., Courville, A., & Schiele, B. (2017). Movie Description. *International Journal of Computer Vision*, 123(1), 94–120. <https://doi.org/10.1007/s11263-016-0987-1>
- Salway, A., Tomadaki, E., Vassiliou, A. (2004). Building and analysing a corpus of audio description scripts. Department of Computing, University of Surrey. Retrieved from https://andrewsalway.work/wp-content/uploads/2020/02/tiwo_television_in_words_deliverable_2-1.pdf
- Sirés, P. M. (2016b). Making the inaccessible accessible: a panorama of audiovisual translation, accessibility and audio description in Japan. *Asiadémica*, 07, 24–50. <https://dialnet.unirioja.es/descarga/articulo/5294998.pdf>
- Slobin, D. I. (1987). Thinking for speaking. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 13, 435. <https://doi.org/10.3765/bls.v13i0.1826>
- Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömquist & L. Verhoeven (Eds.), *Relating events in narrative*, Vol. 2. *Typological and contextual perspectives* (pp. 219–257). Lawrence Erlbaum Associates Publishers.
- Snyder, J. (2007). Audio Description: The Visual Made Verbal. In *The International Journal of the Arts in Society*. Retrieved from http://www.audiodescribe.com/about/articles/ad_international_journal_07.pdf
- Snyder, J. (Ed.). (2010). Audio description guidelines and best practices (Version 3.1). *The American Council of the Blind*. Retrieved from https://specialviewportal.ru/uploads/docs/docs_26.pdf
- Soler Gallego, S. (2019). Defining subjectivity in visual art audio description. *Meta*, 64(3), 708–733. <https://doi.org/10.7202/1070536ar>
- Sonesson, G. (2014). Translation as a double act of communication. A perspective from the semiotics of culture. In Y. Wang & H. Ji (eds.), *Our world: A Kaleidoscopic network*,

proceedings of the 11th world congress of semiotics of IASS in Nanjing, October 5–9, 2012, Vol. 3, 83–102. Nanjing: Hohai University Press.

- Stampoulidis, G. (2019). Stories of resistance in Greek street Art: A cognitive-semiotic approach. *The Public Journal of Semiotics*, 8(2), 29–48. <https://doi.org/10.37693/pjos.2018.8.19872>
- Şulha, P. (2023). Audio describing emotions from a semiotic perspective: Mucize Doktor. *RumeliDE Dil Ve Edebiyat Araştırmaları Dergisi:/RumeliDe Dil Ve Edebiyat Araştırmaları Dergisi*, 33, 1376–1388. <https://doi.org/10.29000/rumelide.1283494>
- Szarkowska, A. (2011). Text-to-speech audio description: towards wider availability of AD. *The Journal of Specialised Translation*, 15, 142–162.
- Tair, S. A., Haider, A. S., & Alkhaldeh, M. (2023). Verbalizing visual characterizations of race in the audio description of *Netflix*. *British Journal of Visual Impairment*. <https://doi.org/10.1177/02646196231203594>
- Tor-Carroggio, I. & Vercauteren, G. (2020). When East Meets West: A comparison of audio description guidelines in China and Europe. *Hikma*, 19(1), 167–186. <https://doi.org/10.21071/hikma.v19i1.12197>
- Trudgill, P. (1995). *Sociolinguistics: An Introduction to language and society*. <https://ci.nii.ac.jp/ncid/BA26949779>
- Trudgill, P., & Hannah, J. (2013). International English. In *Routledge eBooks*. <https://doi.org/10.4324/9780203785225>
- Udo, J. P., & Fels, D. I. (2009). “Suit the action to the word, the word to the action”: an unconventional approach to describing Shakespeare’s hamlet. *Journal of Visual Impairment & Blindness*, 103(3), 178–183. <https://doi.org/10.1177/0145482x0910300307>
- Uulu, Z. Z. T., Şimşek, S. Ç. S., & Antonova-Ünlü, E. (2021). The effect of audio description on film comprehension of individuals with visual impairment: A case study in Turkey. *British Journal of Visual Impairment*, 41(1), 130–142. <https://doi.org/10.1177/02646196211020058>
- Van Dijk, T. A. (1977). Context and cognition: Knowledge frames and speech act comprehension. *Journal of Pragmatics*, 1(3), 211–231. [https://doi.org/10.1016/0378-2166\(77\)90035-2](https://doi.org/10.1016/0378-2166(77)90035-2)
- Vercauteren, G. (2016). A translational and narratological approach to audio describing narrative characters. *TTR/TTR. Traduction, Terminologie, Rédaction*, 27(2), 71–90. <https://doi.org/10.7202/1037746ar>

- Vercauteren, G., & Orero, P. (2013). Describing Facial Expressions : much more than meets the eye. *Revista De Traducció*, 20(20), 187–199. https://ddd.uab.cat/pub/quaderns/quaderns_a2013n20/quaderns_a2013n20p187.pdf
- Vogindroukas, I., Stankova, M., Chelas, E., & Proedrou, A. (2022). Language and speech characteristics in autism. *Neuropsychiatric Disease and Treatment, Volume 18*, 2367–2377. <https://doi.org/10.2147/ndt.s331987>
- Wardhaugh, R., & Fuller, J. M. (1986). *An introduction to sociolinguistics*. <http://ndl.ethernet.edu.et/bitstream/123456789/34482/1/7.pdf>
- Jakobson, R. (1959). On linguistic aspects of translation. In *Harvard University Press eBooks* (pp. 232–239). <https://doi.org/10.4159/harvard.9780674731615.c18>
- Zlatev, J., Devylder, S., Defina, R., Moskaluk, K., & Andersen, L. B. (2023). Analyzing polysemiosis: language, gesture, and depiction in two cultural practices with sand drawing. *Semiotica*, 2023(253), 81–116. <https://doi.org/10.1515/sem-2022-0102>
- Zlatev, J., Żywicznyński, P., & Wacewicz, S. (2020). Pantomime as the original human-specific communicative system. *Journal of Language Evolution*, 5(2), 156–174. <https://doi.org/10.1093/jole/lzaa006>

APPENDIX 1: PLOT SUMMARIES

Atypical (season 1, episodes 1–7)

Sam, a teenager with autism, navigates the complexities of dating with the help of his therapist Julia and encouragement from his friend Zahid. His mother, Elsa, seeks distraction from family stress by starting an affair with a bartender named Nick. Meanwhile, Sam's sister, Casey, faces her own challenges after an incident at school and explores a new relationship with a technical school student named Evan. As the family deals with personal struggles and revelations, Sam is trying to figure out how to tell that you are in love while looking for his own path amidst the chaos.

Totally Completely Fine (season 1, episodes 1–4)

When Vivian's grandfather passes away in his sleep, she inherits his clifftop home. Soon it turns out that the cliff is regularly visited by people attempting to die by suicide. Vivian learns that her grandad had been talking people out of jumping for years, saving hundreds of lives. After saving the life of runaway bride Amy, Vivian starts to take on the responsibility of helping those who come to the cliff while also struggling to navigate the relationships with her two brothers. Flashbacks to her as a young girl reveal that Vivian blames herself for a car accident that killed her mother. By constantly finding herself in situations where she has to save others, Vivian, who is tormented by guilt and haunted by a feeling of rejection, finds a true friendship and accepts that many things are beyond her control.

APPENDIX 2: AVERAGE SENTENCE LENGTH

The first formula (1) was applied to display the word count per cell, after which the second formula (2) counted the number of full stops in the cell, which amounted to the number of sentences. The only other end-of-sentence punctuation mark in audio description is an ellipsis, which the formula sometimes treated as no-period and other times as three periods. This issue was dealt with manually, to ensure the calculations are correct. Finally, to determine the average number of words per sentence, a third formula (3) was applied, which divided the word count per cell into the number of sentences.

(1) =LEN(CLEAN(Cell))-LEN(SUBSTITUTE(CLEAN(Cell); " "; ""))+1

(2) =LEN(Cell) - LEN(SUBSTITUTE(Cell; ". "; ""))

(3) Cell A/Cell B

APPENDIX 3: LOCAL COHERENCE

ID	Image	Dialogue	Sound	AD	Show
1	One of the men, whose name, as we learn now, is Hendrix.	Lawyer: Now, Hendrix!	x	The man with the mustache.	<i>Atypical</i>
2	Amy's face.	Amy: Aww... Amy: Jesus.	x	Amy.	<i>Totally</i>
3	Elsa opens the freezer filled with boxes.	Sam as narrator: ...in order to be a good boyfriend to Julia, I need a practice girlfriend first.	x	A freezer full of hashbrowns.	<i>Atypical</i>
4	Sam sits with his back straight.	Sam: I don't like the feeling of the seat on my back, so I sit like this.	x	Straight-backed.	<i>Atypical</i>
5	Vivian lies on the bed. Vivian sits up.	x	[sound of glass breaking]	By the sound of breaking glass, Vivian sits up.	<i>Totally</i>
6	Dane makes a circle around his face.	Dane: And just to be clear. This... Dane: Only horror and utter confusion.	x	His face.	<i>Totally</i>
7		Vivian: There's only one bed, so...	x		<i>Totally</i>

	Vivian gestures away.			Vivian gestures away.	
8	<p>Casey runs. She cuts across the driveway.</p> <p>Casey stops and screams.</p> <p>She looks over her right shoulder. A neighbor calmly places garbage in a bin. Casey continues running.</p>	x	[scream]	<p>Casey takes an evening jog through a suburban neighborhood. She cuts across the driveway before stopping in the middle of the street.</p> <p>x</p> <p>She looks over her right shoulder. A neighbor calmly places garbage in a bin. Casey continues running.</p>	<i>Atypical</i>
9	Casey in the school hallway in a crowd of students.	<p>Bailey: Oh no. I'm sure whoever did this just feel terrible.</p> <p>Casey: <u>Bailey</u>.</p>		<p>A girl in white stands nearby.</p> <p>Casey taps the girl in white's shoulder.</p> <p><u>The girl in white, Bailey,</u> turns and Casey punches her in the face.</p>	<i>Atypical</i>
10	Outside her house, Casey sees Evan, who offers her a Snickers bar.	<p>Evan: I brought you <u>this</u>.</p> <p>Casey: What is up with your family bringing me sweets?</p> <p>Evan: I don't know. I like you.</p> <p>Casey: Why?</p> <p>Evan: 'Cause most people don't stand up to assholes, but you do.</p>		<p>x</p> <p>He offers her <u>the Snickers</u>.</p>	<i>Atypical</i>

11	Sam talking to Madison.	Madison: You're really weird. And <u>you wear the same clothes every day.</u>		Now, in Julia's office, <u>Sam wears the same outfit.</u>	<i>Atypical</i>
12	John and Alejandro at home.	Alejandro: <u>Here?</u>		Alejandro stands beside a tattered office chair. John considers <u>it.</u>	<i>Totally</i>

APPENDIX 4: FLASHBACKS

Show	Dialogue	Image	AD
<i>Atypical</i>	Sam: I don't like the feeling of the seat on my back, so I sit like this.	Sam on a city bus. The city landscape changes to that of Antarctica, Sam's passion.	Next, Sam on a city bus. Straight-backed. A female passenger observes Sam with furrowed brows. She gives a tight smile and looks away quickly. A large iceberg dominates the horizon. Penguins waddle across the ice. Sam draws a Penguin.
	Sam (narrator): In a natural setting, a rooster will entice hens into mating by spreading his feathers and performing a demonstrative dance. Elsa: Jeez!	Elsa in the kitchen. She smiles as she remembers talking to the bartender earlier. Sam is in his room, but his voice continues narrating while the scene in the kitchen plays.	Elsa grates a lemon peel. She and the bartender talk. Sam's dad grabs her from behind.
	Zahid: Step one. Be hella charming. Sam: Julia, did you know that when Roald Amundsen trekked Antarctica, he would feed his weaker sled dogs to the stronger ones so they could travel more efficiently? He made them cannibalize each other! Julia: What? Sam: What's your favorite meal?	Sam and Zahid at the store (Techropolis). Montage with Sam in Julia's office.	Then, Sam and Zahid sit behind the counter at Techropolis. xxx Julia raises her eyebrows as her mouth falls open. Sam takes notes.
	xxx	The stop sign turns into the word "SLUT" in Elsa's mind.	They arrive at an intersection and Elsa stops the car. She looks up at the stop sign, which reads: "SLUT". She leans forward to

			look closer and grips the wheel tightly.
	xxx	Elsa's husband face turns into the bartender's one in Elsa's mind.	He approaches her. She wraps her arms around his neck and kisses him. For a moment, his face turns into Nick's. She recoils. She looks away quickly before glancing back at him and smiling.
	<p>Woman: What about you, Elsa? Have you done any self-care lately?</p> <p>Elsa: No. Same old, same old.</p> <p>Elsa: I've been... bowling.</p> <p>Woman: Bowling?</p> <p>Elsa: It's crazy, right? But it's so much fun!</p>	<p>A group of moms icing cupcakes in Elsa's kitchen before school party.</p> <p>Elsa flashes to her cheating on her husband with the bartender.</p> <p>Back from the flashback.</p> <p>Flashback.</p> <p>Back from the flashback.</p> <p>Flashback.</p>	<p>Elsa straddles Nick and kisses him as she unbuttons his shirt.</p> <p>She ices a cupcake.</p> <p>She lies on a bed with Nick on top of her.</p> <p>Elsa in bed with Nick.</p>
	<p>Sherice: I'm so sorry. Casey's my best friend and...</p> <p>when I heard you were leaving, I just got so upset...</p>	<p>Sherice and Elsa in a car.</p> <p>Sherice and Casey in a room.</p>	xxx
<i>Totally</i>	xxx	<p>Flashback.</p> <p>Back from the flashback.</p>	<p>Images flash of the young girl in a car.</p> <p>Back to the woman.</p>

xxx	Flashback. Back from the flashback. Flashback.	Now, faded images: flash of sunlight through treetops. A young Vivian looks out a back window of a moving car. She grabs the headrest of the passenger seat and tugs on it. Her mother looks at her angrily while driving. Flash to the present: Vivian stirs in bed. In flashback, Vivian's mother swerves the car to the right.
xxx	Flashback. Back from the flashback.	Flash to Hendrix and Vivian dancing. Back to the living room with Amy.
xxx	Vivian at the cliff. Flashback. Back from the flashback. Flashback. Back from the flashback. Flashback. Back from the flashback.	In the moonlight, she arrives at the cliff and then stops. With her arms crossed, she looks out. Flashback to her as a young girl, in the backseat of a moving car. She walks her fingers across the window. Young Vivian watches her mother's eyes in the rearview mirror. Back to adult Vivian looking out over the water. Then back to her mother's eyes in the rearview mirror getting droopy and closing. Her mother wakes and sits up as a car passes. Back to Vivian in the moonlight, her eyes closed. Then back to her mother falling asleep as the wheels are swerving into the truck. In the moonlight, Vivian opens her eyes and stares off.
xxx	Flashback. Back from the flashback.	In flashback, Vivian's mother steers into an oncoming track. Back to the mediator.

APPENDIX 5: REPETITION IN AD

4.1 Repetition used in AD to reflect repetition in the visual content

(1) They move out of view on the right, leaving only the beach house in sight. They reappear, pushing the bins from right to left. They disappear from view, leaving only the beach house. They reappear, pushing the bins from left to right. (*Totally*)

(2) A message from Nick reads “U up?” ...After a moment, she takes her phone out again and rereads Nick's message, “U up?” ...Nick’s message “U up?” remains unanswered. (*Atypical*)

(3) Sam attempts to smile, but really just shows his teeth. Sam repeats the toothy look. He hands Julia his notes. He begins to write. Toothy look again. (*Atypical*)

(4) First, she types: “Thank you for your continued interest”, but then she grimaces before deleting the message. She starts to type again, smiling as she does, and the message reads: “Let's get naked”. She deletes that message as well, as Sam enters the room. (*Atypical*)

(5) She stares at it for a moment before looking around to see if anyone is watching... After she is finished, she looks around again, and then turns and walks away. (*Atypical*)

(6)

Images flash of a young girl.

Images flash of the young girl in a car.

Flash to a little girl, seated in the back seat of a moving car. She plays with a slinky toy.

Flash to the face of the young girl.

An image flashes of the little girl in the back seat with the slinky toy.

An image flashes of the little girl in the car.

Images flash of the little girl in the car throwing the slinky toy in the front row of the truck.

Images flash of the young girl in the back seat of the car. She tosses a water bottle with a woman in the front, driving.

Now, faded images: flash of sunlight through treetops. A young Vivian looks out a back window of a moving car. She grabs the headrest of the passenger seat and tugs on it. Her mother looks at her angrily while driving. Flash to the present: Vivian stirs in bed. In flashback, Vivian's mother swerves the car to the right.

Flashback to the car swerving. A man in the passenger seat looks back at little Vivian.

A flashback shows Vivian as a young girl in the back seat of a moving vehicle. She plays with dinosaur toys. Her mother smiles as she drives. She looks in the rearview mirror. Young Vivian smiles back at her. As they drive, sunlight flashes through treetops.

In flashback, Vivian's mother steers into an oncoming truck. (*Totally*)

4.2 Repetition used to narrate prominent traits of the characters

SHOW	CHARACTER	TRAIT	REPETITION IN AD
<i>Atypical</i>	Sam	Hypersensitivity	<p>Now in a school hallway Sam walks with his <u>headphones on</u>.</p> <p>Sam wears large <u>headphones</u>.</p> <p>Sam rounds the corner wearing his <u>headphones</u>. He approaches them. He removes his <u>headphones</u>.</p> <p>Now, at school, Sam sits on the hallway floor wearing <u>headphones</u>.</p> <p>She hesitates for a few seconds before tapping him on the shoulder and pantomiming her removing the <u>headphones</u>. He removes the <u>headphones</u>... He immediately puts his <u>headphones</u> back on...</p> <p>Sam <u>sits in a small alcove underneath the desk</u>.</p> <p><u>Headphones</u> on, reading a textbook.</p> <p>Sam <u>hesitates</u> as Zahid holds out his arms.</p> <p>Then, in his room, Sam sits alone on his bed. His <u>hood pulled tight around his head</u>.</p> <p>Sam sits still for a moment before reaching out to remove his <u>hood</u>.</p> <p>Now, Sam lies in bed with a <u>blanket wrapped tightly around him</u>.</p> <p>Sam's <u>tightly pulled hood frames his face</u>.</p>
<i>Atypical</i>	Sam	Interest in Antarctic wildlife	<p>Sam stares at the <u>penguins</u>.</p> <p>As she continues, she finds a <u>diagram of a penguin</u>, in which its skeleton is displayed as if it were an X-ray image. Then, she sees a drawing of a continent</p>

			<p>bearing the legend “<u>Antarctica</u>”. At the center of the land mass, a label reads “<u>South Pole</u>”.</p> <p>Now, at Newton High, Sam stands in front of his locker. <u>Antarctic wildlife pictures</u> hang inside of his door.</p> <p>A large <u>iceberg</u> dominates the horizon. <u>Penguins</u> waddle across the ice.</p> <p>Sam draws a <u>penguin</u>.</p> <p>He holds up a slender silver chain. Hanging from the chain is a black and silver pendant in the shape of a <u>penguin</u>.</p>
<i>Atypical</i>	Elsa	Covering her mouth, suppressing emotions	<p>She <u>covers her mouth</u> as she chews.</p> <p>She <u>holds a clenched fist to her mouth</u>.</p> <p><u>Her hands jumps up to cover her mouth</u>, before she turns and walks towards the door, <u>covering her face</u>.</p>
<i>Totally</i>	Vivian	Alcohol abuse	<p>Now, she looks at a console record player and <u>drinks</u> from a bottle of alcohol.</p> <p>She <u>drinks</u> a bottle of beer as she sits at the coffee table and plays solitaire.</p> <p>She <u>drinks</u> from a bottle of hard alcohol as she enters a cluttered room.</p> <p>She <u>drinks</u> from the bottle and slumps into a chair in the office.</p> <p>She <u>drinks</u> from a bottle of clear alcohol.</p> <p>Amy appears in the doorway with papers. Vivian <u>drinks</u>.</p> <p>She <u>drinks</u> from a beer can.</p> <p>Nearby, Vivian pours a <u>drink</u>.</p> <p>Vivian collects <u>liquor bottles</u>.</p> <p>Now in the dark, Vivian <u>drinks</u> from a bottle.</p>

<i>Totally</i>	Vivian	Lost in thought, disoriented	<p>Vivian sits at the end of the bed. She <u>looks off</u>.</p> <p>Vivian <u>stares off</u>.</p> <p>Vivian stares toward Dane. She looks towards the kitchen. She sits back against the tub and <u>stares off</u>.</p> <p>Vivian looks down, then <u>glances off</u>.</p> <p>She <u>stares blankly ahead</u>.</p> <p>Vivian <u>stares blankly</u>.</p> <p>She <u>appears lost in thought</u>.</p> <p>Vivian <u>appears lost in thought</u>.</p>
<i>Totally</i>	Hendrix	Dieting, always has a soft drink to replace food	<p>Hendrix stares as he <u>sucks on a drink straw</u>.</p> <p>He has a <u>takeout drink</u> and a golden urn.</p> <p>He <u>sucks on his drink straw</u> and hands Vivian the urn.</p> <p>...squeezes a <u>Tetra Pak of strawberry milk</u>.</p>
<i>Totally</i>	Amy	Spiritually inclined	<p>Amy <u>extends her arms and opens her hands</u>.</p> <p>Amy <u>smiles and spreads open her fingers</u>.</p> <p>Amy approaches and <u>spreads out her fingers</u>.</p>