



**Università
degli Studi
di Palermo**

AREA RICERCA E
INNOVAZIONE
SETTORE DOTTORATI E



DIPARTIMENTO DI SCIENZE UMANISTICHE

DOTTORATO DI RICERCA IN
STUDI UMANISTICI

CICLO XXXVII

***New Perspectives on Audio Description:
Enhancing creative accessibility for didactic purposes***

SSD: ANGL-01/C

Dottoranda

Maria Luisa Pensabene

Coordinatrice

Prof.ssa Marina Castiglione

Tutor

Prof.ssa Valentina Castagna

Co-Tutor

Prof.ssa Alessandra Rizzo

Anno Conseguitamento: 2026

To my family...

Acknowledgements

My sincerest thanks go to my PhD supervisor, Professor Valentina Castagna, for her dedication to the highest scholarly standards and for the intellectual honesty she encouraged me to maintain. The depth of her academic experience and her constructive insights have been fundamental for the development of this thesis.

I would also like to express my deepest gratitude to my co-supervisor, Professor Alessandra Rizzo, the prime mover behind this path who believed in me from the very start. Her unique vision and fresh perspectives have been an incredible source of inspiration to me. I am truly thankful for her sharp focus and unwavering commitment.

Just as importantly, I owe a debt of gratitude to both my supervisor and co-supervisor for their empathy and personal support during the most demanding stages of my PhD.

I also wish to thank my coordinator, Professor Marina Castiglione, to whom I extend my sincere appreciation for her guidance and, above all, her patience.

I would also like to thank the Department of Humanities at the University of Palermo for access to equipment and essential resources that made my research possible.

A huge thank you goes to Dr Gabriele Uzzo from the University of Palermo and all the team at SudTitles s.r.l. Throughout our years of working together, they have offered a unique environment to exchange ideas on the technical and cinematic demands of Audio Description. I am deeply grateful for their expertise, which helped me refine my professional skills, and for the genuine friendship that made my six-month placement with them so rewarding.

Last, but absolutely not least, this thesis is a reflection of the unwavering support I received from my family. I cannot thank them enough for their steadfast encouragement and for being a constant presence that provided the strength I needed to persevere through these years.

Table of Contents

Table of Contents.....	6
List of Acronyms.....	9
List of Figures.....	12
List of Tables.....	13
Preface.....	14
Abstract.....	15
Introduction.....	16
 1 Chapter 1 – Accessibility Studies and Media Accessibility.....	 20
1.1 From Audiovisual Translation to Audiovisual Accessibility.....	20
1.2 The Medical and Social Model of Disability.....	23
1.3 Accessibility Studies.....	25
1.4 Media Accessibility.....	29
1.5 Accessible Filmmaking: Access to Content and Creation.....	31
 2 Chapter 2 – Audio Description. From Accessibility to Language Education...	 36
2.1 Audio Description: Origins and Status of the Discipline.....	36
2.2 A general overview of Vision Impairment.....	40
2.2.1 A Clinical Perspective on Blindness and its Perceptions.....	42
2.2.2 Vision beyond vision: exploring the unseen.....	44
2.3 Audio Description as a complex process beyond human rights.....	46
2.4 An overview of existing guidelines.....	48
2.5 Audio Description in Reception Studies.....	50
2.6 The Creative Turn.....	51
2.7 Audio Description in Foreign Language Education.....	53
 3 Chapter 3 – Didactic Audio Description in Foreign Language Education: A Practical Test.....	 59
3.1 Project Design: Context and Aims.....	60
3.2 Audiovisual Genres: Three Different Proposals.....	62
3.3 A functionalist approach to text types.....	64
3.4 <i>Billy Elliot</i>	69
3.4.1 Exploring the film’s Themes and Meanings.....	71
3.4.2 Text composition of <i>Billy Elliot</i>	73
3.5 <i>The Ellen DeGeneres Show</i>	76
3.5.1 Exploring <i>The Ellen DeGeneres Show</i> ’s Themes and Meanings.....	76
3.5.2 Text composition of <i>The Ellen DeGeneres Show</i>	79
3.6 <i>Aagaswadi – Village in the Sky</i>	82
3.6.1 Exploring the documentary’s Themes and Meanings.....	82
3.6.2 Text composition of <i>Aagaswadi – Village in the Sky</i>	84
3.7 Learners: the participants in the didactic experiment.....	86
3.8 CADAD as a type of Didactic Audio Description.....	87
3.9 CADAD Task: Design and Implementation.....	92
3.9.1 Introductory Phase.....	93

3.9.2 <i>Billy Elliot</i> : PP1, SWP1, PTP1.....	98
3.9.3 <i>The Ellen DeGeneres Show</i> : PP2, SWP2, PTP2.....	100
3.9.4 <i>Aagaswadi – Village in the Sky</i> : PP3, SWP3, PTP3.....	103
3.10 <i>Frazier® - VideoToVoice</i> as a Pedagogical tool for CADAD.....	103
 4 Chapter 4 – A corpus-based Analysis: Assessing CADAD through <i>AntConc</i> and ADLAB Guidelines	108
4.1 Corpus Analysis: <i>AntConc</i>	108
4.2 Quantitative Analysis: measuring lexical frequency in CADAD scripts.....	111
4.3 Assessing priorities in lexical patterns: an ADLAB-informed analysis.....	149
4.4 Analysis of <i>Billy Elliot</i> CADAD: enhancing creativity through emotion.....	153
4.4.1 Characters, action, language and style.....	154
4.4.2 Spatio-Temporal Dimension.....	157
4.4.3 Figurative Language.....	159
4.4.4 Film Language.....	160
4.5 Analysis of <i>The Ellen DeGeneres Show</i> CADAD: creativity through action and physical description.....	161
4.5.1 Characters, action, language and style.....	161
4.5.2 Spatio-Temporal Dimension and Film Language.....	165
4.6 Analysis of <i>Aagaswadi – Village in the Sky</i> CADAD.....	166
4.6.1 Characters, action, language and style.....	166
4.6.2 Spatial Dimension and Film Language.....	171
4.7 <i>Frazier®</i> – “Markers” and “Snippets”: Frequency and Application.....	171
4.8 Culture-Specific References	173
4.9 Framing the emerging priorities in the CADAD scripts.....	178
4.10 Evaluation process: learners’ language proficiency through AD scripts.....	180
 5 Chapter 5 – Qualitative Research Methodology	184
5.1 Analyzing Functional Adequacy and Creative Modulation of AD Scripts.....	185
5.2 Overall Discussion.....	191
5.3 The Architecture of Subjectivity: Appraisal Theory.....	193
5.3.1 Theoretical Framework.....	194
5.3.2 Analysis of Audio Description Segments through the Attitude System.....	200
5.3.3 Overall Patterns of Subjectivity in the Scripts.....	206
 Concluding Remarks	208
Bibliography.....	214
Sitography.....	251
Filmography.....	253

List of Acronyms

AAO: American Academy of Ophthalmology
ACB: American Council of the Blind.
AD: Audio Description.
AdjP: Adjective Phrase
ADP: Audio Description Project.
ADTs: Audio Description Techniques.
AE: Attitudinal Epithet (used in the Appraisal Theory analysis).
AENOR: Asociación Española de Normalización y Certificación (Spanish Association for Standardization and Certification).
AFM: Accessible Filmmaking.
AoA: Action-Oriented Approach.
ARDELE: *Audiodescripción como Recurso Didáctico en la Enseñanza del Español como Lengua Extranjera* (Audio Description as a Didactic Resource in Teaching Spanish as a Foreign Language).
ARSAD: Advanced Research Seminar on Audio Description.
AS: Accessibility Studies.
AVA: Audiovisual Accessibility.
AVST: Audiovisual Source Text.
AVT: Audiovisual Translation.
AVTS: Audiovisual Translation Studies.
AVTT: Audiovisual Target Text.
B/VIP: Blind or Visually Impaired Persons.
BE: *Billy Elliot*.
BFI: British Film Institute.
BVI: Blind and Visually Impaired.
C: Classifier (used in the Appraisal Theory analysis).
CADAD: Creative Didactic Audio Description.
CEFR: Common European Framework Reference for Languages.
CSRs: Culture-Specific References.
DAD: Didactic Audio Description.
DAT: Didactic Audiovisual Translation.
DFC: Didactic Free Commentary.
DMC: Digital Multimodal Composing.
EAA: European Accessibility Act.
EACEA: Education, Audiovisual and Culture Executive Agency
EBU: European Blind Union.
ES: *The Ellen DeGeneres Show*.
FLE: Foreign Language Education.
FLL: Foreign Language Learning
FTII: Film & Television Institute of India.
HCI: Human-computer interaction

IAPB: International Agency for the Prevention of Blindness.
 ICD-10: International Classification of Diseases.
 ICEVI: International Council for Education of People with Visual Impairment.
 ICFDH: International Classification of Functioning, Disability and Health.
 ICTs: Information and Communication Technologies.
 IP: Introductory Phase.
 KWIC: KeyWord In Context.
 L1: First Language / Native Language.
 L2: Second Language / Foreign Language.
 LE: Language Education
 MA: Media Accessibility.
 MAS: Media Accessibility Studies.
 MIM: Ministero dell'Istruzione e del Merito (Ministry of Education and Merit).
 MLVs: Multiple Language Versions.
 MUR: Ministry of University and Research.
 NFB: National Federation of the Blind.
 NG: Nominal Group (used in the Appraisal Theory analysis).
 NIPAI: New International Performing Arts Institute
 NLP: No Light Perception.
 NRRP: National Recovery and Resilience Plan.
 P: Process (used in the Appraisal Theory analysis).
 PBS: Public Broadcasting Service.
 PCTO: *Percorsi per le Competenze Trasversali e per l'Orientamento* (Pathways for Transversal Skills and Orientation).
 POUR: Perceivable, Operable, Understandable, and Robust.
 PP1, PP2, PP3: Pre-task Phase(s).
 PP: Prepositional Phrase (used in the Appraisal Theory analysis).
 PSL: People with Sight Loss.
 PTP1, PTP2, PTP3: Post-task Phase(s).
 RA: Relational Attribute (used in the Appraisal Theory analysis).
 RNIB: Royal National Institute of Blind People.
 RP: Relational Process (used in the Appraisal Theory analysis).
 RS: Reception Studies.
 SDH: Subtitling for the Deaf and Hard of Hearing.
 SFG: Systemic Functional Grammar.
 ST: Source Text.
 SWP1, SWP2, SWP3: Script-Writing Phase(s).
 TAD: Traditional Audio Description.
 TBLT: Task-Based Language Teaching.
 TS: Translation Studies.
 TT: Target Text.
 TTS: Text-to-Speech.
 UD: Universal Design.
 UICI: *Unione Italiana dei Ciechi e degli Ipovedenti*.

UNCRPD: United Nations Convention on the Rights of Persons with Disabilities.
UPIAS: Union of Physically Impaired Against Segregation.
UX: User Experience.
VI: Vision Impairment.
VIPs: Visually Impaired Persons.
VITS: *Aagaswadi Village in the Sky*.
VoD: Video-on-Demand.
VtV: *VideoToVoice*.
WBU: World Blind Union.
WCAG: Web Content Accessibility Guidelines.
WCWB: World Council for the Welfare of the Blind.
WHO: World Health Organization.
WYSIWYS: What you see is what you say.
ZPD: Zone of Proximal Development.

List of Figures

- Figure 1. MA at the intersection between AVT and AS, resulting in AV).
- Figure 2. The formation process of AS according to Greco (2019).
- Figure 3: Particularist-account design vs UD of an urban environment (AI-generated).
- Figure 4: Particularist ICT environment vs UD ICT environment (AI-generated).
- Figure 5. The battle for *Les Tourelles* from “*The Messenger: The story of Joan of Arc*” (Luc Besson, 1999).
- Figure 6: Didactic AD diagram adapted from Talaván *et al.* (2024).
- Figure 7: Posters and factsheets of the three audiovisual genres.
- Figure 8: Billy goes through emotional turmoil (source: mabumbe.com)
- Figure 9: The Clint Eastwood costume moment (source: langhingsquid.com)
- Figure 10: A frame from *VITS* (source: alteff2020.org).
- Figure 11: Percentages of the participants’ knowledge of AVT and AD.
- Figure 12: Author’s elaboration of the use of CADAD in the task.
- Figure 13: *Frazier*’s main interface showing described segments and the editable text box (source: videotovoice.com).
- Figures 14 and 15: Examples of descriptions for *Billy Elliot* and *The Ellen DeGeneres Show*.
- Figure 16: *AntConc* 4.3.1 main interface (source: laurenceanthony.net).
- Figure 17: ADLAB guidelines’ mental model creation process in story-reconstruction (2015).
- Figure 18: Distribution of adjective use across semantic categories.
- Figure 19: Marker insertion during the Ellen script writing task on *Frazier*®.
- Figure 20: The three metafunctions through AD samples and their lexicogrammatical realizations (author’s elaboration).
- Figure 21: The Appraisal Systems (adapted from Martin and Rose, 2003; cf. Rizzo and Pensabene, 2021).
- Figure 22: The Attitude system and its three sub-systems based on Martin and Rose (2003 and adapted by Lipson (2006).

List of Tables

- Table 1: Author's integration of functional characteristics of text types according to Reiss (1971/2000) and Mazur (2020).
- Table 2: Functional classification of audiovisual text types and implications for macro-level audio description strategies (author's elaboration).
- Table 3: Functional characteristics of The Acceptance Letter Scene (author's elaboration).
- Table 4: Descriptive dimensions and associated challenges in the Billy Elliot acceptance letter scene.
- Table 5: Functional characteristics of the Halloween Costumes excerpt (author's elaboration).
- Table 6: Descriptive dimensions and associated challenges in the Halloween Costumes segment.
- Table 7: Functional characteristics of *Aagaswadi* - Village in the Sky's opening scene (author's elaboration).
- Table 8: Descriptive dimensions and associated challenges in the opening sequences of *Aagaswadi* - *Village in the Sky*.
- Table 9: Participation details by school, including duration of the activity and daily attendance.
- Table 10: Structure of the CADAD workshop.
- Table 11: The CADAD scripts that make up the corpus.
- Table 12: Incomplete and missing descriptions across the three genres.
- Table 13: Implicit and explicit references to spatial dimension.
- Table 14: Examples of *Frazier* bookmarks.
- Table 15: Culture-specific elements in the three genres.
- Tables 16 and 17: Percentage breakdown of the most recurring word types in the three CADAD sets.
- Tables 18 and 19: Key differences between B2 and C1 proficiency levels in mediating a text and creative writing production (adapted from CEFR/CV, 2020).
- Table 20: Appraisal-based Subjectivity continuum.
- Table 21: Overview of subjectivity patterns across the three genres.

Preface

This work is really the result of three different strands in my life: my years as an English teacher, my time working professionally as an audio describer, and a curiosity that goes all the way back to my childhood.

Growing up, I occasionally spent some time with my great-uncle Valentino. He had lost his sight to diabetic retinopathy and this made me wonder what it was like for him to live in a world without sight, especially if faces and objects that were once visible now existed only in his mind's eye. I can still see him pulling animals out of my toy farm set, trying to guess what they were by their shape and describing them back to me in his own way. Those moments were definitely my first introduction to tactile and verbal representation.

Later, when I became more acquainted with audiovisual translation during my university years, it felt like everything had finally come together, so I decided to take the plunge into the theoretical and professional intricacies of audio description.

Looking back, I don't think it could have turned out any other way.

Abstract

Over the last few decades, the multi-faceted nature of Audio Description (AD) has transformed this modality of Audiovisual Translation (AVT) into a versatile resource that benefits sighted audiences as well. This evolution has pushed scholarly attention past the traditional boundaries of accessibility and into more experimental landscapes, giving rise to a prolific literature that investigates the creative and emotive dimensions of AD, as well as its complex discursive layers. Most recently, there has been a surge of interest in its didactic value within Foreign Language Education (FLE), where AD is being used as a practical, complementary tool to foster learners' language proficiency.

Although AVT has long been a staple of FLE pedagogy, it has recently gained momentum as a supplementary pedagogical methodology following a new elevated vision of the learner suggested by the CEFR/CV (Common European Framework Reference for Languages, 2001, 2018, 2020), where the language student is presented as a "social agent" acting in real life situations and carrying out multiple tasks. This Action-Oriented Approach (AoA) allows learners to put their skills to the test through active involvement, while also improving their intercultural competence. This work explores the potential of Creative Didactic Audio Description (CADAD) based on a practical experiment conducted at the University of Palermo in 2024, involving two groups of Italian upper-secondary school learners of English who produced AD scripts for three diverse audiovisual genres: drama (*Billy Elliot*), talk show (*The Ellen DeGeneres Show*), and observational documentary (*Aagaswadi – Village in the Sky*). The study combines quantitative and qualitative methods. First, the scripts generated through the writing tasks are analyzed using *AntConc* to examine lexical dispersion and identify the most salient lexical patterns within the corpus. The analysis then adopts a qualitative perspective, interpreting the scripts along a creativity continuum to classify them according to Bardini's Audio Description Techniques (ADTs) and Rizzo and Spinzi's strategies, while also evaluating their functional adequacy. Finally, the Appraisal Theory is layered in to account for evaluative and subjective stances in the learners' language. To deconstruct the "architecture of subjectivity" in the scripts, the Attitude sub-system (Affect, Judgement, and Appreciation) of Appraisal is applied to evaluate the learners' creative agency and their intersubjective relation with the AD elements.

Introduction

For decades, the media landscape and accessibility endured what Pablo Romero-Fresco (2019) calls the "long divorce." Until the late 2000s, audiovisual modes such as Audio Description (henceforth AD) were considered as an afterthought in the production process, that is a *post-hoc* intervention added at the end of a project. Following the development of a new paradigm of societal accessibility from the mid-2010s onwards (see Greco, 2016), these two areas have been redefined within the broader landscape of Audiovisual Accessibility (AVA), which considers access as a proactive resource that belongs at the very heart of the creative process. Expanding from this theoretical framework, this study aims to validate AD as a leading creative force that has pushed the boundaries of Media Accessibility (MA), evolving from a tool for the blind and visually impaired into a fertile ground for experimental and innovative uses beyond accessibility, also entailing the spheres of didactics and foreign language teaching.

Starting from the core idea that descriptions are a creative art form in their own right, this five-chapter thesis focuses on filmic AD and explores what happens when we take AD out of its accessibility landscape and bring it into language learning environments to make the most of its pedagogical potential.

The **first chapter** establishes the theoretical and historical framework of the research, beginning with an overview of the "accessibility revolution" and its impact on various disciplines. It explores how these fields have undergone a three-fold paradigm shift, moving from particularist accounts of accessibility to universalist ones, from maker-centered to user-centered design, and from reactive to proactive methodologies. This transformation, caused by the ideological tension between the medical model of disability and the emergence of a social model, is particularly evident in MA. Now part of the broader, interdisciplinary field of Accessibility Studies (AS), this research area identifies access as a fundamental human right and a pre-condition for social participation.

It is here that the concept of Accessible Filmmaking (AFM) surfaces as a proactive resolution to the "long divorce" between filmmaking and the needs of diverse audiences, calling on the integration of translation and accessibility from the *ex-ante* stages of production. In particular, the discourse surrounding AFM highlights how this participatory approach to AD has brought about a plethora of new creative methodologies.

The **second chapter** revolves around the evolutionary path of AD as an academic discipline and a professional practice, starting from its origins up to its experimental applications that go well beyond the purpose of making content accessible. After introducing the status of the discipline and its creative advancements, I examine visual impairment from a clinical perspective, suggesting that – other than its objective constraints – blindness can be a valuable resource for creative inspiration. The chapter then goes on to dissect the core tenets of filmic AD, the most significant studies to test audience reception and the existing guidelines, before narrowing the focus on creativity. Ultimately, this discussion broadens out to recognize AD as an integration tool for sighted users, including its application in Foreign Language Education (FLE) to enhance language proficiency, and especially writing production skills.

The **third chapter** is concerned with a practical test that I conducted at the University of Palermo with two groups of Italian upper-secondary school learners of English. The participants, who had an average B2 level of proficiency, engaged in a creative didactic AD task with three different audiovisual genres, namely drama, talk show, and a non-fictional documentary. Consequently, the audiovisual source texts (AVST) are categorized from a Translation Studies (TS) perspective to analyze their textual composition and identify: a) the main intersemiotic challenges and b) the most suitable descriptive strategies for didactic purposes. The test span over a two-day workshop on AD, during which the learners were asked to produce three AD scripts for each genre using *Frazier® - VideoToVoice*, an innovative cloud-based platform for script-writing. In particular, the task pushed the participants to prioritize creative expression over objectivity according to the specific demands of each genre; therefore, the use of vivid adjectives, adverbs, verbs, as well as figurative language was highly encouraged, prompting them to find subjective and interpretive solutions.

Moving forward, the **fourth chapter** introduces the analytical framework, focusing on a software-aided quantitative investigation of the emerging lexical priorities from the CADAD scripts and a more detailed observation of the learners' word choices inspired by the ADLAB guidelines (2015). The final section delineates a task-based assessment framework, adapting CEFR/CV (2020) B2-C1 descriptors to intralingual AD as a mediation strategy and a creative production method.

Finally, the **fifth chapter** pivots away from a didactic context to adopt a qualitative research perspective: the main purpose here is to understand the subjective "value" learners placed

on the images they were describing. To do so, I selected a number of segments to assess the functional adequacy of the strategies used for “translating” the most salient visual cues into verbal renditions: their work is then analyzed following Bardini’s (2020) and Rizzo and Spinzi’s (2023) taxonomy of AD techniques (ADTs) along a “creativity continuum”. This analytical framework is further enriched by the application of the Attitude sub-system of the Appraisal Theory, namely an extension of M.A.K Halliday’s Systemic Functional Grammar (SFG) model. This complementary methodology is used to identify how learners interpret audiovisual elements and express subjectivity through their evaluative stances. The analysis helps determine how these linguistic resources are embedded within grammatical structures such as attitudinal epithets, relational processes, and nominal groups to single out the presence, type, and intensity of the learners' appraisal.

Chapter 1 – Accessibility Studies and Media Accessibility

This introductory chapter begins by laying out the theoretical and historical dimensions of the research, and looks at how MA has grown from a sub-area of AVT into the broader field of Audiovisual Accessibility (AVA). The discussion then goes on to examine the shift from the medical model of disability to a social one that redefines accessibility as a fundamental human right and not simply a remedial intervention. From there, the chapter outlines the core principles of AS focusing on the proactive approach as a way to finally resolve the “long divorce” (Romero-Fresco, 2019) between filmmaking and accessibility. This leads into an analysis of AFM where the chapter makes the case for bringing translation, disability, and access requirements into the heart of the creative process.

Ultimately, the focus narrows to the creative potential of AD showing how, over the last decade, the practice has evolved into a tool for innovation that goes well beyond the basic goal of making content accessible.

1.1 From Audiovisual Translation to Audiovisual Accessibility

This work revolves around the ever-changing nature of AD, specifically how its evolution has opened up new creative and pedagogical perspectives. However, we cannot truly venture into this terrain without first taking a step back to place it within the broader framework of MA and, more generally, accessibility. We may argue that, if AD serves as a bridge, accessibility provides the solid foundation on which it rests. Before becoming a creative practice, the description of images - and, more broadly, audiovisual content - functions primarily as a means of making media accessible to a wider range of audiences, particularly blind and visually impaired people (B/VIP) (Walczak & Fryer, 2017:7). This affirmation resonates with a widespread understanding of social and cultural inclusion, far from a particularist account of accessibility and much closer to a universalist one. The implication behind it is that accessibility is no longer a specific concern for people with disabilities, but an enabling principle for all human beings (Stephanidis & Emiliani, 1999), a vision that aligns with the core tenets advocated by AS. Over the last decade, this area of research has been critically investigating accessibility phenomena and methodologies that have contributed to the paradigm shift from a reactive to a proactive design of our society.

This major overhaul is visible in various fields and, to a larger extent, in MA, the sub-area that has been growing exponentially within AVT (Díaz Cintas, 2005; Orero, 2005). MA refers to “a set of theories, practices, services, technologies and instruments providing access to audiovisual media content for people that cannot, or cannot properly, access that content in its original form” (Szarkowska *et al.*, 2013). Originally born to break down sensory barriers – with subtitling for the deaf and hard of hearing (SDH) and AD as the main modes of access – MA has gradually extended its scope to linguistic barriers (Díaz Cintas, 2005; Orero & Matamala, 2007). More recently, it has been reconceptualized in wider, universalist terms, encompassing access to media products, services, and environments for all individuals (Greco, 2016b). Traditionally, MA has been positioned within AVT which, in turn, falls under the wider umbrella of Translation Studies (TS) and differs from it on the nature of its linguistic output and the techniques it puts into practice.

Sometimes referred to as “screen translation” “film translation”, or “multimodal translation” (cf. Chaume, 2019; Gambier, 2003; Zanotti, 2022), it involves the localization of audiovisual media content through different translation practices (Díaz Cintas, 2019). What truly sets AVT apart from traditional translation is its multimodal configuration, where meaning is spread across various channels. Consequently, the process is best understood as a form of multimodal translation that goes far beyond a simple search for linguistic equivalence and cannot be relegated to the label of adaptation. It is therefore more accurate to speak of “transadaptation” (Chaume, 2018; Gambier, 2003; Neves, 2005), as this term best captures the complex intermodal changes at stake.

AVT has gained momentum over the last few decades. Interestingly, Díaz Cintas (2019) goes back to the 1995 centenary of cinema as the symbolic catalyst for both scholarly and professional interest in a field that has since seen an unprecedented surge in subtitled, dubbed, and audio described content. On the one hand, this shift has led to a proliferation of audiovisual content; on the other, it has sparked deeper investigations into the theoretical complexities and foundations of the discipline. However, the emergence of digital technology and global streaming platforms (e.g., Netflix and Amazon Prime to name but a couple) has changed the way media content is consumed, to the extent that the traditional divide between dubbing and subtitling countries now appears obsolete (Dore, 2018; cf. Gambier, 2003; Chaume, 2013). Furthermore, cultural and technological developments, as well as the demand for simultaneous delivery, has also prompted a “retranslation” of audiovisual material (Dore, 2018; Chaume, 2018).

Nevertheless, the blurred lines between MA and AVT have been the subject of much contention among scholars, as the two have been understood either as strongly convergent, overlapping disciplines (cf. Remael, 2022; Díaz-Cintas, 2022) or as autonomous domains (Greco, 2018; Greco & Jankowska, 2022), the latter of which is the most widespread position today. Yet, between these two poles lies the bold assertion that MA almost functions as a superordinate category that goes beyond translation itself (Gambier, 2004). As Greco (2018:218) argues, “it is [...] an interdisciplinary area that criss-crosses many well-established fields, including TS and AVT. However, it cannot be [...] exclusively reduced to any of them because its true nature is that of being an area within the field of AS”. This position is particularly compelling as it mirrors the ontological shift we are seeing today. Building on the development of a new paradigm of societal accessibility, we can assume that MA is the nexus where the (inter)semiotic modes of AVT meet the ethical mandates of AS to build inclusive media, and if we accept this, it is therefore more accurate to speak instead of AVA (cf. Uzzo, 2024).

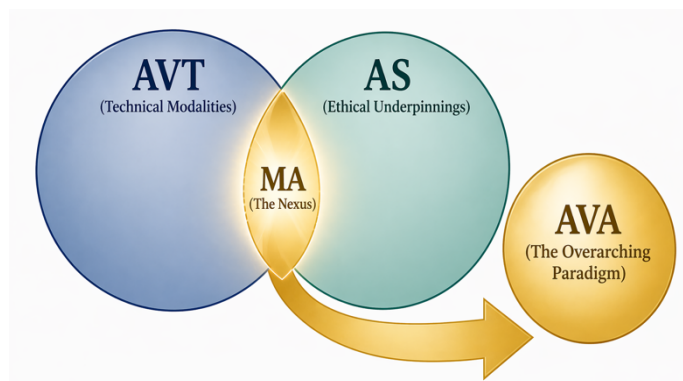


Figure 1: MA at the intersection between AVT and AS, resulting in AVA.

In this light, disability becomes part of a wider conversation about diversity, and the traditional notion of access gives way to a more inclusive vision shaped by Universal Design (UD) (Di Giovanni, 2024; cf. Gosset *et al.*, 2009). To make sense of this transformation, we must first consider how the idea of disability itself has changed over time.

1.2 The Medical and Social Model of Disability

The accessibility revolution began to unfold as disability was rethought. This transformation was not abrupt, but the culmination of a long process of critique and reformulation. It is a legacy dating back to the 1970s, born from the ideological tension between over-medicalized accounts and the need for a social model of disability, promoted by British organizations like the *Union of Physically Impaired Against Segregation* (UPIAS) and the *Liberation Network of People with Disabilities*.

Today, a third model of disability seems to be prevalent worldwide and is often depicted in the media. It is the moral model, one that sees disability in symbolic terms, as either a punishment for character flaws or as a triumph of human survival to complex experiences. However, I will focus the scope of this work on the clash between the medical and social frameworks, as it is within the two that the concept of disability is redefined.

Historically, the term “disabled” was predominantly used to describe individuals perceived as having an “intrinsic defect, a chronic illness, or disease” that put them in a position of “functional limitation” and in a state of “subnormal functioning” (Cabrera, 2009; Roulstone, 2016). Therefore, “disabled” people were dependent upon others, an assumption that shaped that ableist vision based on the relationship “between the active able-bodied helper and the passive disabled object of attention” (Brechtin *et al.*, 1981). Conversely, the social model places that faultiness on society itself rather than the individual. First of all, the social model makes a clear distinction between impairment and disability. According to the *International Classification of Impairments, Disabilities, and Handicaps* (ICIDH) proposed by the World Health Organization (WHO, 1980:47), the former concerns “any loss or abnormality of psychological, physiological, or anatomical structure or function”, while the latter refers to “any restriction or lack (resulting from an impairment) of ability to perform an activity [...] in the range considered normal for a human being”. Shakespeare (2010) argues that it is the outside world that disables physically impaired people and disability is something imposed on top of their impairments causing them to be excluded from full participation in society.

The social model is essentially a social construct that defines disability as a conflict between a person’s impairment and a society that refuses to accommodate them; the main concern is not so much to “cure” the individual as it is to remove barriers and pass anti-discrimination laws. Its main strength is that it stops seeing people as inherently “broken”

and starts seeing them as “an oppressed group”. This change is more than philosophical and did not come without its blind spots, though. For example, one of the weaknesses of the social model is that it risks brushing aside the fact that an impairment is a problem and, “rather than opposing medicalization, it can be interpreted as rejecting [...] prevention [...] and rehabilitation” (*ibidem*). An additional drawback could be the concept of a barrier-free utopia, that is an environment where all imposed barriers are removed, that is, *de facto*, a difficult realization despite the valuable approaches of the UD, as some resource constraints make it impossible to overcome every barrier (Finkelstein, 1981). This is where the legal framework intervenes to bridge the gap between ideology and rights. As Greco (2022) notes, contemporary claims for accessibility are to be sought within the *United Nations Convention on the Rights of Persons with Disabilities* (UNCRPD). Crucially, the aim of the UNCRPD was not to create new rights from scratch, but to ensure that existing human rights were finally made effective for persons with disabilities (Broderick, 2020). By using the Convention as a backbone, the discourse then moves from “charity” to “justice” and the accessibility issue is no longer a favor or an afterthought, but rather an essential requirement for equal access to environments, services that the law already demands.

However, the most sophisticated evolution of the medical-social dichotomy is, perhaps, not seeing disability as a lack of something, but as “a way of perceiving and orienting to the world” (Titchkosky, 2011), which, in turn, calls for the biopsychosocial model advanced by the World Health Organization (WHO, 2001) through the *International Classification of Functioning, Disability and Health* (ICFDH). As seen in recent pedagogical insights (Tait & Silveira, 2023), this model frames disability as an interaction between biological traits, psychological wellbeing, and social environments, where the priority is understanding the specific needs of the individual without losing sight of the collective commitment to adapting the environment. Most importantly, this perspective resonates with the need for a new critical pedagogy of accessibility (Greco, 2019) as a social and cultural practice.

Turning now to the audiovisual sphere, if disability is a different way of “orienting” to media, the role of MA changes as well. From a universalist point of view, it is about making multimodal content accessible for all and designing inclusive services that can be used regardless of sensory or physical disabilities (Di Giovanni *et al.*, 2022). Therefore, if those barriers are reimagined as catalysts for diversity rather than difference, we know that MA is no longer about translating visual or audible cues for those who cannot see or hear;

instead, it is more about validating the diversity of perceptual experiences, which is the idea behind a proactive design of accessibility.

1.3 Accessibility Studies

Today, accessibility is more crucial than ever. It has become a fundamental part of social and cultural belonging that pervades a vast range of fields, sparking a wealth of new topics and innovations (Greco, 2019). In the context of MA, interdisciplinarity is particularly evident. AVT scholars have contributed to expertise in subtitling, AD, and dubbing, thus conversing with linguists who have analyzed language use and multimodal meaning-making, as well as with scholars in disability studies who have provided theoretical frameworks centered on inclusion, agency, and the social model of disability. Today interdisciplinarity also relies on human-computer interaction (HCI) and computer science experts who have developed assistive technologies and AI-driven accessibility tools, together with psychology and cognitive science scholars who have investigated how users perceive, process, and remember accessible content. Last but not least, education and pedagogy explore how accessibility can be taught and integrated into professional training. Besides, interdisciplinarity also extends to professional practice and producing accessible content often involves collaboration among translators, audio describers, software developers, filmmakers, educators, disability advocates and end users. Increasingly, accessibility follows a user-centered or participatory approach, in which disabled people are involved throughout the design, production, and evaluation process. This collaborative model recognizes that accessibility cannot be achieved through technical expertise alone but requires the combined contributions of different disciplines and stakeholders.

In a word, interdisciplinarity has triggered an accessibility revolution culminating in the emergence of AS, a research field “concerned with the investigation of accessibility processes and phenomena, and [...] accessibility-oriented methodologies” (Greco, 2018:205).

This revolution is essentially the result of two other major changes: on the one hand, the cultural revolution produced by human rights which established access as a prerequisite for human dignity; on the other, the information revolution, since Information and Communication Technologies (ICTs) “are changing the world we live in, showing how

the ‘digital divide’ reinforces existing social inequalities, *i.e.*, those who have access and those who don’t” (Uzzo, 2024).

Rather than delving into the disciplinary status of the field or the controversies surrounding its potential misapplication, this section aims to outline its core purposes and its practical contributions to MA. Over the years, accessibility has been the center of attention of scholarly research, finally evolving as an autonomous discipline, comprising its own specific topics, models and methods (*ibidem*; cf. Greco, 2019) (Fig. 3):

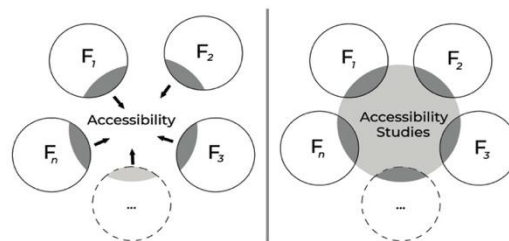


Figure 2: The formation process of AS according to Greco (2019).

Most importantly, accessibility has had a massive impact across a vast array of fields to the extent that these fields have been living through a three-fold paradigm shift: from a particularist account to a universalist account, from a maker-centered to a user-centered approach, and last, but not least, a change from a reactive to a proactive approach.

The particularist account sets the limits of accessibility in a barrier-free environment for specific groups of people with particular needs. By contrast, the universalist approach widens the scope of accessibility to all human beings, advocating for an inclusive society for everyone through a UD. This all-embracing approach involves the implementation of services and environments that “meet the needs of the broadest possible range of users, regardless of age and ability” (Clarkson & Coleman, 2015). Therefore, ramps are not only intended for wheelchairs, but also for parents carrying strollers, as well as workers transporting goods (Pujiyanti & Ikaputra 2025). And the same goes for digital and computational technologies, whose accessibility requirements must follow and maintain the POUR rule from the get-go, that is they must be *perceivable*, *operable*, *understandable* and *robust* (Web Content Accessibility Guidelines [WCAG], 2023) (Figs. 3 and 4).



Figure 3: Particularist-account design vs UD of an urban environment (AI-generated).



Figure 4: Particularist ICT environment vs UD ICT environment. (AI-generated).

If accessibility is to be framed as a non-negotiable human standard, then the needs of all end-users are all of equal importance. This principle is closely related to another paradigm shift: the move from a maker-centered to a user-centered approach, which redefines users as "bearers of knowledge for the investigation of accessibility processes and phenomena" (Greco, 2018). In this model, users hold a seat at the table during the

entire implementation process, so the maker's knowledge is not the only perspective that matters; they can no longer be locked into their own point of view, as they will always be biased by their own technical expertise. As Greco (*ibidem*) notes, a major consequence of the traditional maker-centered approach has been a series of disconnects between stakeholders, most notably the maker-user gap.

To explain this, I will draw on an example from usability experts Nielsen and Loranger (2006) regarding the web-purchase experience. They identify three widening tiers of epistemic distance – and the resulting failures of intuition – between the creator and the consumer. I will call the first the "designer-as-user" tier, which occurs when both the designer and user share the same technical background. For example, creators of open-source software often build products that succeed only as long as the audience remains a group of tech-obsessed enthusiasts. Their familiarity with the system's internal logic creates a gap with non-expert audiences that can only be bridged by external validation. Similarly, a "proximal tier" occurs when a designer uses the product themselves and falls victim to their false consensus bias. By assuming their personal experience is a universal baseline, they ignore the diverse mental models of the wider public. Accordingly, this results in an environment that is technically functional but cognitively inaccessible to those who do not share that specific knowledge (e.g., standard e-commerce websites). Lastly, I call the third tier the "absolute experiential divide", which happens when the creator has zero experience of the user's reality, as is the case for sighted designers placing accessibility buttons at the bottom of a page for visual balance without ever testing them with a blind user.

Another significant gap identified by Greco (2018) is the maker-expert-user gap, a "triangular disconnect" that occurs when the maker, the expert, and the user speak different languages. Specifically, makers tend to focus on technical feasibility, whereas experts usually look to theoretical standards, and users are the ultimate judges of lived experience. Closing this gap requires recognizing that the user's experiential knowledge is just as relevant as the maker's expertise and must be integrated into the design. Consequently, the implementation of an artefact follows a breakdown into a series of *ex-ante*, *in itinere*, and *ex-post* stages. This ensures accessibility is not just an afterthought, but a core part of the production process from its earliest stages (Romero-Fresco, 2019). This approach calls upon the final shift: the move from a reactive to a proactive account, according to which accessibility is a foundational blueprint and not a remedial intervention

occurring in response to failure. In other words, adopting a proactive mindset entails integrating access into the initial *ex-ante* stages, thereby redirecting attention from fixing mistakes to preventing barriers and ensuring equal access for all.

1.4 Media Accessibility

This accessibility revolution has been equally visible in the field of MA and has reframed the discourse surrounding it. From a theoretical point of view, the process has been longer and more complex though. As already discussed, MA was originally conceived as nothing but a sub-domain of AVT, limited to sensory barriers. So, accessibility of media content was upheld if subtitles for SDH and AD for the blind were provided. For example, the first edition of *The Arts and 504: a 504 Handbook for Accessible Arts Programming* (1985) insists on “captions” for the SDH and “verbal descriptions” for the B/VIPs (Greco & Jankowska, 2020). Conversely, Romero-Fresco (2018) hints at a “wider notion of MA” that blows the doors off the traditional definition towards a universalist perspective, suggesting that access to media (here intended as audiovisual material) is a fundamental right for everyone. However, as long as the stigma persists, accessibility remains shackled to the medical model of disability, when, on the contrary, a universalist approach demands that we break out of it. The shift comes after the realization that, other than an impairment, sometimes disability is also situational and may occur to any person at any given moment. For example, an elderly person or a student trying to read in a library may experience access issues due to small font sizes or loud background noises, not to mention a non-native speaker watching a foreign film. And, it is exactly this latter scenario that widens the scope of MA as a universal necessity at the intersection with AVT. If we think that MA is limited to sensory barriers, we are ignoring the reality that disability in the digital age is a functional mismatch other than a medical diagnosis. The term “disabled” may refer to anyone (be they sensory-impaired, a linguistic minority, or a non-native speaker) who is denied the ability to access and enjoy entertainment services, which, in turn, is a reminder that linguistic barriers are just another form of access issue. This is what happens in the current media landscape, in which intermodal and intersemiotic processes have become part of the user’s daily digital experience. Take social networks, for example: users on

Facebook or *TikTok* are constantly exposed to interlingual respeaking, subtitles, and audio subtitling, whether intentionally or not.

Arguably, much of the changes also depend on the fact that audiences are more in control of their entertainment experience today than they were ten years ago, a new reality that is impacting AVT and MA with respect to producers, practitioners, audiences and workflows (Zanotti & Sánchez-Mompeán, 2025). The global success of services such as Netflix, Amazon Prime and Disney+ among others, has spurred a growing demand for localized media, while also increasing the availability of audiovisual content translated into multiple languages through captioning and revoicing techniques (*ibidem*). Nonetheless, the fact that the paucity of resources often precludes translating content into a wide range of languages (Tuominen *et al.*, 2023) is a separate issue that falls beyond the scope of this work.

What is worth bringing into focus, instead, is the fact that a universalist approach in MA is not a theoretical preference but a direct response to contemporary media consumption increasing user agency. However, we cannot deny that a maker-user gap endures because of production companies, which still see these features as external additions. This is the case of AD, for example; descriptions are often remembered as post-production solutions, stretching the void between filmmaking and accessibility, or what Romero-Fresco (2018) calls “the long divorce” (2018), a disconnect dating back to the “talkie” revolution of the late 1920s¹. Following the advent of synchronized sound, translation and accessibility – once handled in-house – ended up being outsourced to distribution companies as an *ex-post* addition. There seems to be every reason to believe that filmmakers and accessibility practitioners rarely spoke to each other, thus moving the accessible version of a film away from the director’s original vision (Branson, 2024; Romero-Fresco, 2013).

It is at this point that the concept of Accessible Filmmaking (AFM) comes to the surface as a proactive resolution to the long-debated split, calling on the integration of accessibility, [...] “budgeted for from the beginning of the project” [...] “through the collaboration between the creative team and the accessibility experts” (Romero-Fresco, 2018:192; Branson, 2024:15)

¹ The “Talkie Revolution” refers to the industry-wide transition from silent film to synchronized sound following the success of *The Jazz Singer* (directed by Alan Crossland, 1927).

1.5 Accessible Filmmaking: Access to Content and Creation

AFM represents a fundamental pillar of a proactive approach to MA and a significant departure from the traditional idea of a passive audience, advocating instead for their active participation in the creative process. As discussed in the previous section, historically, the advent of the “talkies” carries a deeper meaning than that of a new semiotic dimension of cinema, namely sound. Silent films were well suited to consumption in many cultural contexts, which was due less to their status as a universal language of images than to their intertitles and the flexibility they provided (Betz, 2001). Flexibility, in this sense, refers to a target-oriented transcreation process through which cultural references and character details are transposed to meet the specific expectations of each national audience. However, as soon as films and other audiovisual content became audible, it was clear that film distribution would involve more than simply replacing intertitles, even though the first clumsy attempts at dubbing and subtitling did not live up to expectations. If such practices appeared rudimentary, then multiple language versions (MLVs) (Vincendeau, 1999) functioned as a more reasonable workaround, before interlingual dubbing and subtitling rose to prominence as more cost-effective procedures in the 1930s. These translational practices certainly facilitated a broader dissemination of media content, but we cannot deny that such audiovisual advancement came at a price, meaning that translation was pushed back to the post-production phase. Over time, this late-stage relegation has built up an afterthought mentality as standard industry practice, the same mentality that leaves sensory-diverse audiences in the blind spots of the media landscape. Building on this, AFM means that accessibility – whether for foreign-language speakers or deaf and blind viewers – is an act of creative responsibility and placing the international audience on the same level as the sensory-impaired audience is the way forward to an accessible (and successful) filmmaking model. As noted by Romero-Fresco (2019), both interlingual translators and accessibility professionals are creative collaborators who should be included in the production pipeline from its earliest stages. Their early involvement is paramount to preserving the director’s creative vision, ensuring a level of aesthetic and functional equivalence that would be otherwise lost. More specifically, the term “equivalence” here refers not so much to the pursuit of a formal correspondence in the replication of textual meaning (Vinay & Darbelnet, 1995; Nida & Taber, 1969/1982) as to a comparable

experience between an audience watching a subtitled or audio-described version and the original intended audience.

Currently, a growing number of mainstream directors, *i.e.*, John Sayles, Jim Jarmusch, Danny Boyle, and James Cameron are gravitating toward this model, while others have extended this approach further by drafting their own tenets of creative sovereignty for distributors, as is the case with Quentin Tarantino and Alejandro González Iñárritu (Romero-Fresco, 2017). However, these examples remain the exception rather than the rule, as the majority of distributors are still somewhat reluctant to surrender the final word on localized versions. Conversely, such resistance has made independent cinema the ideal environment for developing AFM methodologies – whether focused on language or sensory barriers; indeed, films like Michael Chanan’s *Secret City* (2012), Alastair Cole’s *The Colours of the Alphabet* (2016), or Pete Middleton and James Spinney’s widely recognized *Notes on Blindness* (2016) are considered as the milestones of a proactive and user-centered approach to MA. The latter – an adaptation of Australian theologian John Hull’s audio diaries about losing his sight in a botched cataract operation – is notable for integrating accessibility features, “such as audio description and enhanced soundscapes, directly into the production process rather than adding them as an afterthought” (Gibbs, 2016; Romero-Fresco, 2017).

To this end, independent filmmakers often experiment with creative subtitling, using idiosyncratic fonts and effects to transform a functional, linguistic tool into a core aesthetic choice (Foerster, 2010). An example of this is Edgar Pêra’s (2011) Portuguese neo-Gothic drama *O barão* (*The Baron*), where “expressive and kinetic” subtitles are used to move the viewer from immersion to aesthetic amazement (Kapsaskis, 2024). More importantly, Pêra designed the subtitles as a graphic element of the film, which is a reminder of the core tenets of AFM: subtitles and AD belong in the *ex-ante* production stage and must not be added as a remedial intervention. When accessibility is baked into the creative process, then the director is able to preserve their authorial intent regardless of the language, while also maintaining the aesthetic and affective impact of the film. However, that is a lot easier in theory than in practice, as mainstream cinema is often stifled by copyright rules and hierarchies. For this reason, independent documentaries have emerged as a more fertile ground for AFM, for their “realistic and observational life-caught-unawares approach” borne out of a “[...] relationship of trust between the filmmaker and the participants that makes the filmmaker’s involvement [...] an issue of ethics and basic responsibility. On the

other hand, more innovative [...] documentaries are often based on a sense of creativity that can open the door to new approaches to translation and accessibility” (Romero-Fresco, 2017). Regarding this, London-based company *Archer’s Mark* has turned this principle into their calling card, working alongside the *British Film Institute* (BFI) to develop the *Accessible Filmmaking Guide* (Romero-Fresco & Fryer, 2019). Interestingly, even a giant like Netflix has caught on to the fact that their bottom line depends on translated content, meaning that accessibility is not just a courtesy, but a way to promote “otherwise unseen opportunities for international growth” (Miller, 2023).

The most ambitious side of AFM is that it is not just about *watching* films, but *making* them for the sake of participatory accessibility. At this point, we are referring to a freer notion of access to creation, where people with disabilities take on the mantle of storytellers, as is the case with Dangerfield’s work with deafblind persons acting as co-directors. Indeed, Romero-Fresco (2019) makes a critical distinction between “Disability Art” and “Disability in the Arts”, arguing that films like *Joining the Dots* dealing with disability as a theme are a great step forward, but the ultimate goal for AFM is the active participation or representation of people with disability in the arts, regardless of the themes.

All in all, these AFM methodologies are reminiscent of the so-called notion of non-cinema (Brown, 2016b), in which the “non” suggests a deviation from the standard cinema of Hollywood. In fact, AFM works are deliberately imperfect, challenging the viewer to participate in a reality that is often neglected by mainstream cinema. With specific reference to AD, AFM rejects the *ocularcentric* bias of high-budget productions in favor of a more haptic and sensory-led approach (Elsaesser & Hagener, 2010) that is more consistent with the true nature of AD. The theatre offers significant insights in this regard, particularly in how Traditional Audio Description (TAD) – a practice designed to be inclusive – becomes exclusive by presenting information through a strictly *ocularcentric* viewpoint, leaving blind audiences reliant on choices made by sighted viewers (Fryer, 2018; Cavallo, 2015; Thompson, 2015). This new wave of thinking has paved the way for alternative and creative approaches to AD, most notably the “flipping descriptions” method, that consists of “taking the description out of the hands of the sighted person and handing it to the audiences it was designed to support” (Romero-Fresco & Dangerfield, 2022; cf. Hayhoe, 2018). One interesting example of this was Geering (2019)’s “Rationale Method” for describing dance moves through the participatory creation of blind audiences, who provided beatboxing sounds to represent their own sense of the pulse and pace of the movements.

There is no question that these practices mark a considerable shift toward participatory accessibility, but making them work on a large scale is still a major hurdle. Indeed, they represent a significant move toward a more proactive and authentic user experience (UX), yet the reality of integrating participatory models into both live performance and AFM remains a daunting task.

To some degree, we have seen signs of progress with the emergence of specialized roles, such as the accessibility manager (Remael *et al.*, 2019) and the director of accessibility and translation (Fryer, 2021); however, there is still a long way to go, as these practitioners have certainly set the benchmark for inclusive media, but the rest of the industry has been slow to follow suit so far. Currently, these figures are more prevalent in live performance than in cinema, and even there, they tend to exist as pioneering roles within niche or experimental projects rather than as permanent fixtures.

Arguably, accessibility adds an additional layer of intricacy to an already complex creative process (*ibidem*), and even though this on-and-off presence may feel like a tokenistic choice (LaPointe, 2023), more often than not, it comes down to budgetary constraints and an afterthought mentality that is proving hard to shake. That being said, a full discussion of these issues lies beyond the scope of this study; what matters most for the present work is how, over the last decade, AD has become a creative force in itself, leading to experimental and innovative uses that go way beyond accessibility.

Chapter 2 – Audio Description. From Accessibility to Language Education

This chapter follows the evolution of AD from an accessibility service to a creative and pedagogical discipline. It first gives a brief overview of its institutional context, by comparing the legislative frameworks in Italy, the US and the UK and then moves on to the clinical and sociological realities of vision impairment. In this framework, blindness is not understood as a creative dimension in its own right, but as a condition that fosters the development of creative possibilities for cinematic perception.

The chapter proceeds to examine the tension within the “Accessibility Turn”; here the traditional objective model clashes with more interpretive approaches that bring emotional engagement and the creative role of the describer to the fore. Finally, the focus shifts to the classroom, where the pedagogical potential of AD in FLE is analyzed; building on this, the final part of the chapter explores the framework of Didactic Audiovisual Translation (DAT), demonstrating how AD has recently become a powerful instrument for foreign language acquisition.

2.1 Audio Description: Origins and Status of the Discipline

Over the past two decades, the heightened profile and maturation of MA have sparked a renewed scholarly interest in AD, now recognized as a unique AVT modality that mediates the intersemiotic relationship between the visual and the verbal. AD owes much of its consolidation to biennial conferences like *Media for All*, an influential forum for debate on major research developments beyond translation alone, where strong emphasis also falls on production, distribution, and audience engagement (Perego, 2014). It is in this context that filmic AD has risen to prominence as one of the core pillars of MA, together with SDH: always straddling competing academic definitions, it has been described either as a translational activity (Hyks, 2005) or a quintessential manifestation of AVT for its intersemiotic dimension (Gambier, 2004; Perego, 2005; Perego & Taylor, 2012). This is indeed one of the most distinctive traits of AD, which finds its direct parallel in Jakobson’s seminal taxonomy of translation practices (1959/2012); his categorization identifies intersemiotic translation as the interpretation of verbal signs through non-verbal ones, although in the case of AD, this process is actually reversed. (Bardini, 2020). More

specifically, Braun (2008) also defines it in terms of intermodal mediation, focusing on the transfer from the visual mode to verbal formulations, a procedure that relies on the agency of the (audio) describer, that is the audiovisual translator who executes the intermodal transfer. However, it would be reductive to simply refer to describers as accessibility providers: their role is much more complex and requires considerable mastery of film language and image semiotics, as highlighted by Díaz-Cintas (2007). For this reason, academic attention has shifted toward defining describers as creative practitioners engaged in a form of cinematic *ekphrasis*² (Pujol & Orero, 2007; Klishevich, 2023), as well as cultural mediators, whose involvement intensifies in direct proportion to the distance between the source and target cultures (Sanz-Moreno, 2017).

For a long time, AD was considered as a marginal branch of AVT within the broader field of TS, before finally securing its place as an independent discipline of its own right (Taylor & Perego, 2022) with a dedicated conference. Formally known as ARSAD (Advanced Research Seminar on Audio Description), this flagship scholarly event was first organized by the *TransMedia Catalonia* Research Group in 2007 and, over the years, it has evolved into “a full-fledged, two-day conference with over 100 attendees” (Greco, 2022).

If the volume of research on AD has continued to grow exponentially, so too has the plethora of definitions about its theoretical foundations. Nevertheless, while introducing a new conceptualization into such a well-established landscape might feel redundant, it is still useful to be clear about the nature of the practice that inspired the present work. To do so, I will extrapolate the definition proposed in my pilot study (2023), which describes AD as an intersemiotic mode of AVT, where an auditory narration of relevant visual elements is provided to make media content accessible to the B/VIPs. Additionally, this “visual-to-verbal rendition includes a large array of multimodal works and may refer to: descriptions of movement, characters and settings during films or TV shows in the pauses between dialogues, an elucidation of actions surrounding sports events or theatrical performances”, as well as explanations of “the visual features of artworks and objects, often with some accompanying history, context or biographical insights about the artist” (Hutchinson, 2020;

² The Greek term *ekphrasis* mainly refers to “a literary figure that provides the graphic and often dramatic description of a painting, a relief or other work of art” (Pujol & Orero, 2007:49), as well as to “the verbal representation of visual representation” (Heffernan, 2015). This explains why the concept has been integrated into film semiotics and shares a connection with AD.

cf. Pensabene, 2023:52). However, for the purposes of this study, I will limit the scope to filmic AD.

What clearly emerges from these definitions is, without a doubt, the main reason AD came into being before finding a foothold in experimental applications, that is to ensure that media content was fully accessible to the BVI. That being said, historically, the earliest forms of rudimentary descriptions date back to the Middle Ages as oral explanations for pilgrims to understand the visual narratives in churches (Orero & Matamala, 2013), although Snyder (n.d., online) finds its genesis in ancient Greece. However, it was not until the advent of silent cinema that exhibitors began hiring narrators to explain films and read the intertitles to audiences.

In some ways, the fact that this work is essentially centered around creativity resonates with the very origins of descriptions because, in truth, AD itself has its roots in acts of creative improvisation ushering in a move toward accessible media. Its origin dates back to the 1930s, when a young Ronald Reagan extemporized play-by-play broadcasts of baseball games as a sports announcer, demonstrating the power of imaginative description. Similarly, in the 1940s, New York City Mayor Fiorello LaGuardia read newspaper comics over WNYC radio during a strike, adding a descriptive commentary to help listeners visualize the images (Snyder, 2022). However, it was not until the 1960s that a formal concept of AD began to take shape in the United States, with the most relevant initiatives growing from direct involvement of the blind community. For instance, Chet Avery, a Maine-born blind theatre enthusiast proposed audio captions for films and educational content, while working at the U.S. Department of Education, before connecting with Dr Margaret Rockwell, a blind expert in Education and her future husband Cody Pfanstiehl, who had expertise in radio-reading services in Washington DC. Dr Rockwell, founder of the *Metropolitan Washington Ear* reading service (1981), began collaborating with the *Arena Stage* in Washington D.C. to provide AD for theatrical performances. Her expertise eventually expanded into radio and museums, and by 1987, she was working with the Public Broadcasting Service (PBS) to bring AD to television. This momentum culminated in the *Telecommunications Act* of 1996, which officially recognized internet services and encouraged the diffusion of new technologies, providing a solid legal and professional framework for AD to flourish.

However, from an academic perspective, the concept of AD is universally attributed to American author and professor Gregory Frazier, so much so that he is acknowledged as

the father of AD. Mostly unaware of the Planstiehl's' parallel work, Frazier explored the subject in his MA thesis at San Francisco State University in 1975, after recognizing the need for simultaneous image description while watching a film with a blind friend and noticing how much visual information was inaccessible (Perego, 2016). Following this line of thought, the scholar saw the practice's burgeoning potential and sought to establish its academic and normative framework, ultimately devoting much of his career to its development and founding the non-profit *Audio Vision SF* (1991). This pioneering venture became a cornerstone of the industry and was notably recognized with an *Emmy Award* in 1990 for its description of the film *Tucker: The Man and His Dream* (1988).

In Europe, the United Kingdom, was the frontrunner in these audiovisual practices, promoting social inclusion and access to media. Early milestones, such as the *Audetel-3 project* (1991), focused on exploring descriptive styles and digital delivery, but the true shift towards institutionalization occurred with the *Communications Act 2003*. This landmark legislation changed the status of AD from an experimental service into a regulatory requirement, mandating a 10% quota for major broadcasters.

Conversely, the Italian landscape followed a more fragmented trajectory. The first AD did not appear until 1991, when television author Sergio D'Ottavi and *Rai Due*'s director Giampaolo Sodano decided to provide a radio narration of a film simultaneously with its television broadcast. Indeed, the enactment of *Law 112/2004* – widely known as *Gasparri Law* – encouraged a 10% increase in AD coverage, but without binding authority for broadcasters, meaning that compliance was voluntary and dependent on their initiatives. As AD provision remained uneven, the responsibility of ensuring accessibility of audiovisual content was taken up by a number of independent projects. Two of these include *Senza Barriere ONLUS* and *ARTIS Project*. The former is a social cooperative based in the Trento area, founded in 1992 by a group of visually impaired people with the aim of creating accessible multimedia content. In 2004, it launched Italy's first and only Audio Film Library for the Blind, and today it is the country's most active distributor of audio described films for B/VIPs (Perego, 2014). The latter, instead, is a company offering AD, SDH and other services for audiences with sensory disabilities, founded by researcher and accessibility expert Vera Arma.

However, over the years, there has been a gradual move toward standardization, as both countries have introduced regulatory frameworks, each following their own approach. In the United Kingdom, the *Media Act 2024* has effectively "closed the accessibility gap"

(Habicht *et al.*, 2024) by extending the long-standing 10% AD quota from television to major Video-on-Demand (VoD) platforms (Ofcom, 2024; cf. RNIB, 2025). On the other hand, Italy is now experiencing what we may define a legislative catch-up thanks to the *European Accessibility Act* (EAA, Directive EU 2019/882), which regulates AD as a European standard, thus establishing a formal requirement that replaces the voluntary effort.

This shift was also visible in the preparations for *Milano Cortina 2026*, where the rollout of the *Audiodescrizione26*³ protocol marked a historic first for the country and proves that AD is not an optional extra, but a mandatory pillar of an inclusive digital society.

2.2 A general overview of Vision Impairment

Over the past decades, the accessibility debate has taken on a socially oriented dimension, with significant implications for MA and, especially, AD. If, on the one hand, AD is about making audiovisual content usable so that all audiences can enjoy and connect with it meaningfully, on the other, its interpretation has often been that of replicating the experience of sighted audiences, reflecting the medical model logic rather than the emancipatory potential of the social one (Greco, 2022; cf. Udo & Fels, 2010b). In the case of blindness, this ableist assumption has often left the BVI on the sidelines as passive recipients rather than active participants in cultural life. At this point, taking a closer look at the blind community can open a door to understanding what they expect from AD and how they perceive it. The literature to date provides important insights on eye-tracking observations (Di Giovanni, 2013; Walczak & Szarkowska, 2012; Kruger, 2012; Szarkowska *et al.*, 2013), reception studies (Chmiel & Mazur, 2011b; Udo *et al.*, 2005; Orero, 2008a) or experimental methodologies on users' engagement (Fryer & Freeman, 2012a, 2012b; Fryer, 2010a; Igareda & Matamala, 2012). Nonetheless, I believe that a brief overview of blindness should come before any other discussion on the preferences and interpretive capacities of BVI. Understanding blindness in this light is a necessary step to fully appreciate the true nature of AD and, in order to do so, a brief discussion from a clinical perspective follows in this and the subsequent section. According to the International Classification of Diseases,

³ *Audiodescrizione26* is an *iPhone* app that made the Milan–Cortina 2026 Olympics accessible by providing audio descriptions, allowing all spectators to fully experience and engage with the events.

(ICD-11, WHO, 2022), blindness is defined by the evaluation of persisting post-treatment and post-correction visual acuity and visual field loss. Visual acuity is the clarity of vision, determined by the sharpness of retinal focus, the sensitivity of the nervous system, and the brain's cognitive processing. It is typically measured by the ability to recognize black symbols of varying sizes on a white background at a set distance. The visual field, instead, refers to the area that the eye can perceive when looking straight ahead (Bestard-Bou & Arias Badia, 2022).

Clinically, vision impairment (VI) is an umbrella term used to describe all levels of reduced vision – from partial sight to total blindness – that cannot be fully corrected by lenses or surgery. This expression has often sparked controversy in accessibility debates when used to refer to total sight loss, reflecting the belief that language carries power and shapes how marginalized groups are perceived. Terms like “visually impaired” are generally considered politically correct, though they imply a subtle sense of weakness or damage. Similarly, the word “blindness,” carries a negative connotation that overlooks the wide spectrum of visual ability, often leading to misunderstanding. Interestingly, US deafblind author and athlete Brian Switzer (2021), an outspoken advocate for the blind community, has coined the term “visually diverse” inspired by the concept of “neurodiverse” to describe the full range of visual experiences. However, the discussion is not without its more radical voices. Following a 1993 resolution by the National Federation of the Blind (NFB), Kenneth Jernigan argued against the use of the so-called “people-first language”, e.g., “person with blindness” in favor of direct language, e.g., blind person. Jernigan compares the evolution of politically correct to the decline of civilization, where form becomes more important than substance, resulting in ludicrous euphemisms. Echoing this sentiment, the NFB asserts that the word “blind” is respectable and accurate enough to describe a physical condition that does not need linguistic cushioning to validate equality (Jernigan, 1993; NFB, 1993).

However, this should not be taken to mean that misuses of the term do not exist. For example, James H. Omvig, former president of NFB, accounts for a number of commonly accepted dictionaries which define blind as “unable or unwilling to perceive or understand”, or even “not based on reason or evidence” (Omvig, 2009). Arguably, such elucidations are forms of ableist language that can perpetuate a discriminatory resonance reminiscent of the “ghetto effect”, regardless of the intentions behind its usage. Interestingly, Omvig (*ibidem*) suggests that a possible reclamation of the term occurs when people with lack of eye sight come to terms with the fact that they are blind and should

learn to refer to themselves as such when their vision has deteriorated to the extent that alternative non visual techniques are necessary to carry out daily activities.

2.2.1 A Clinical Perspective on Blindness and its Perceptions

According to WHO, all individuals, at some point in their lives, will experience an eye condition that will require appropriate care. These conditions range from minor refractive errors – whether functional or age-related – to major ocular diseases. The former typically include refractive errors such as myopia, astigmatism, and presbyopia, although minor visual disturbances may also arise from non-ocular causes, such as migraine aura, vascular issues, or other neurological or psychogenic factors. The latter consists of one or more temporary, progressive or permanent conditions affecting acuity, peripheral vision, contrast sensitivity, light sensitivity and depth perception. Taken together, these parameters determine whether a person is classified as totally blind, as having low vision, or as experiencing other forms of VI. Total blindness refers to a complete lack of light perception, known as *No Light Perception* (NLP) with binocular visual field residue below 3%; conversely, low vision includes both people with a binocular visual field under 10% and those whose binocular residual vision is not more than 1/20, even with lenses (Italian Law 138/2001, Classification and quantification of visual impairments and rules on ophthalmic assessments).

In general, blindness and low vision may be either congenital or acquired later in life as a result of illness, injury, or retinal degeneration, some of which can be prevented or mitigated through access to quality eye care; however, it should be noted that the causes of VI vary significantly across countries depending on “the availability and affordability of eye care services as well as the level of awareness” (WHO, 2026). This diversity is important to consider, as users’ needs may differ widely depending on the origin and nature of their visual condition.

WHO (2019) estimates that at least 2.2 billion people worldwide experience near or distance vision impairment, of whom 36 million are completely blind. The main causes of distance VI include cataracts (94 million cases) and uncorrected refractive errors (88.4 million). Blindness, on the other hand, is often associated with conditions such as age-related macular degeneration (8 million), glaucoma (7.7 million), diabetic retinopathy (3.9

million), and retinitis pigmentosa (1.5 million), one of the most common hereditary retinal diseases. Presbyopia (826 million cases), by contrast, is the most prevalent condition associated with near VI. Leaving aside the latter, most of the aforementioned pathologies require regular monitoring or special treatments, such as pressure-lowering and steroid drops or other supplements for glaucoma or macular degeneration (American Academy of Ophthalmology, AAO, 2025). However, in the case of serious ocular conditions, such medicaments are not and cannot be definitive treatments, meaning that they can affect a person's overall quality of life, placing significant strain on mobility and social interaction. In fact, it cannot be denied that – beyond the medical classification of VI – the life of a blind person is often associated with greater isolation than that of a person with low vision. For this reason, the World Council for the Welfare of the Blind (WCWB) joined forces with the International Federation of the Blind in 1984 to give birth to the World Blind Union (WBU) organization, which unites major national and international groups addressing the needs of the BVI. Following the principles of the UNCRPD, WBU advocates full and equal participation in society, with accessibility as the third strategic priority, behind human rights and representation (Bestard-Bou & Arias Badia, 2022). The International Agency for the Prevention of Blindness (IAPB) takes initiatives such as the “*Love your Eyes Actions*” to eliminate the main causes of avoidable blindness (Murray Mc Gavin, 1999) and secure access to eye health services. Additionally, the International Council for Education of People with Visual Impairment (ICEVI) focuses on complementary areas, including the importance of literacy and educational resources for children and adults.

At the regional level, the European Blind Union (EBU) works to make sure that the rights and needs of the community are built into laws and public policy, while the Royal National Institute of Blind People (RNIB) in the United Kingdom and the *Unione Italiana dei Ciechi e degli Ipovedenti* (UICI) in Italy also contribute to providing practical and emotional support, including a helpline, talking books, and specialist products, such as screen readers, refreshable Braille displays, and tactile maps⁴.

Looking across all these initiatives, it becomes clear that much of their work ultimately aims to alleviate social and emotional isolation, two things that BVI often have to walk a tightrope between. To tell the truth, living with VI is not inherently isolating, but the subtle pressures of social stigma can push individuals toward potential segregation (Dunlop

⁴ These specialist products are formally known as assistive technologies and typhlotechnical aids (from the Greek *typhlós*, meaning blind), which include any device or equipment designed to improve autonomy, communication, and environmental accessibility for B/VIPs.

et al., 2025). As noted by Michalko and Titchkosky (2024), there has been a tendency in Western modern thought to portray the blind as lacking freedom of movement, prioritizing rehabilitation and the technological mastery of impairment, *i.e.*, the development of Braille. These achievements are celebrated – and with good reason – as triumphs, yet, they tend to perpetuate sight-centered perspectives and a narrowed conception of human potential that hinder the social and collective dimension of blindness. Conversely, the key to overcoming stigma is viewing blindness not as an individual lack, but as a distinct form of movement and connection shaped by family and community interactions. It is about recognizing the cultural meanings and narratives surrounding blindness, seeing blind people as active contributors to knowledge and social life, and ultimately challenging the assumption that vision is the default or superior mode of perception (*ibidem*).

2.2.2 Vision beyond vision: exploring the unseen

This study resonates with the idea that, freed from the literal constraints of sight, blindness itself is a mode of inventive perception and, therefore, a valuable source of creative inspiration. Indeed, a blind person feels the world through imagination, touch, sound, and memory that help them transform absence into a fertile terrain for creative exploration. In this sense, blindness is not a limitation, but a way of seeing that goes beyond the eye, where every experience is reimagined and recreated.

Blindness is often described as light deprivation, which may register as a synonym for “darkness” in the mind of a sighted person trying to conjure up this reality. However, this description turns out to be inaccurate. As psychologist Luca Mazzucchelli (2013) points out, the concepts of light and dark are not properly applicable to blindness – unless light is understood as the use of non-visual senses, and darkness as the absence of such sensory input.

When it comes to imagination, a common misconception is to think that the mind of a blind person is necessarily dark. In the case of individuals experiencing vision loss, the transition clearly involves both physical and psychological adjustment, although this does not mean that they cannot come to terms with their loss over time. However, we should learn to think of imagination as an innate intellectual virtue (Marber & Wilson, 2024, *my emphasis*), regardless of sensory limitations; the difference is that a sighted person’s

mental images are mostly visual and often colorful and those who become blind may still retain visual memory. Conversely, someone who is congenitally blind experiences mental imagery through senses other than sight and is more likely to improve semantic and episodic memory, as reported by research on this subject (Pasqualotto *et al.* 2013).

US film critic and radio host Tommy Edison, who is congenitally blind, explains that he has no visual reference for objects, e.g., a baseball (The Tommy Edison Experience, 2012). Instead of describing it visually, he focuses on its physical properties and how it feels or exists in sound and touch – for example, its shape, texture, weight, and the sound it makes when hit or thrown. He emphasizes that, because he has never seen objects, he has no interest in forming visual images of them in his mind (Nair, 2025).

Generally speaking, in human perception, one sense may become dominant over the others – even for sighted people – and is often the one most recalled when describing a situation. For instance, a person walking along a busy city street at night will experience the neon signs and traffic lights, the roar of engines, the chatter of pedestrians or the smell of street food. When they later describe the scene, they might focus first on the sounds and smells rather than the visual details, because those senses were most striking in that moment.

For all the reasons above, darkness and light are abstract concepts in the mind of blind-born people. However, the color dimension still holds particular significance, since blind individuals tend to make associations between colors and objects: for example, green may bring to mind plants or spring; red is often connected to warmth and fire, symbolizing vitality and pride; blue, as in the sea, may give a sense of calm and reserve, *etc.* With reference to this, Tommy Edison provides a witty and insightful take on his perception of color:

Over the years people have tried to explain color to me, and I just don't understand it. I think the best way to show you is to try and explain to someone who's never heard before what the ocean sounds like, or what the birds sound like. And that's where color is for me. [...] And people will try and explain a sense with another sense. 'It's like the way this smells, maybe' [...]. What? So, you're going to explain what color looks like by my nose? [...] When somebody says 'red' to me I don't quite get it, I know that red is fire. [...] When someone says 'you're in the red' that means you're in some trouble, like financial trouble. [...] Blue. The water is blue, cold or ice is blue, the sky is blue. Now, how can the sky and ice be the same thing? That's weird to me. [...] Same color, it means two completely different things. [...] Black is supposed to be all the colors smudged together and white is the

absence of color? So, when I hear black and white, to me they're just opposites, 'cause one has everything and the other has none. And then there are things that don't have color, like water. It doesn't have color but the ocean does? I don't get that. Colors are hard. How do sighted people keep 'em all straight? (Tommy Edison, 2012)

Going back to AD, chromatic variety is not excluded from the adjectival sub-groups used to describe physical appearance and descriptions of colors are always encouraged and appreciated both in museum and filmic AD, as even people who are congenitally blind know the cultural associations with them (Giansante, 2015:8; cf. Perego, 2018), not to mention that they help to recreate textual links to support cohesion (Taylor *et al.*, 2015). Even Snyder (2014), in his account of audio describers' skills, suggests including colors when editing and culling the most important information from what they see. And, on this very point, he reminds us – once again – that AD comes from a full understanding of what blindness is all about to give language and perception their full depth.

2.3 Audio Description as a complex process beyond human rights

As a productive field of research, AD has attracted a growing number of scholars to explore its untapped potential, turning what was exclusively meant for VIPs (Frazier, 1975) into an AVT modality within MA for all audiences (be they sighted or visually impaired). To some extent, the practice has struggled to secure its own socio-political legitimacy within the broader discourse on accessibility, and paradoxically, this tension stems from the UNCRPD, where AD was first codified as a human right.

However, its evolution can be fully understood in relation to the "Accessibility Turn" in TS. Under the medical model of disability, AD has often been relegated to the role of a compensatory add-on for a functional limitation, that is to say, a niche service providing only the bare essential of a visual narrative for VIPs. Instead, the universalist approach looks beyond this mentality toward a more user-centered paradigm. This has led scholars and stakeholders to call for a reconsideration of the discipline's ontological status within the Convention, resulting in a reconceptualization of AD not as a human right in itself but as a tool for exercising the right to cultural participation (Greco, 2022), as well as a valuable resource for anyone seeking a fuller experience of media content.

From a theoretical perspective and beyond its legal institutionalization, AD sits at the crossroads of MA and AVT, as a form of intersemiotic translation. Yet, this definition captures only part of the picture, since AD goes well beyond a simple transfer from images to words. It is a form of multisemiotic translation (Fryer, 2016; cf. Mazur, 2022) or, more accurately, an aural-verbal narration that covers the original aural-verbal (speech), non-verbal-aural (music), visual-verbal (subtitles and intertitles) and visual non-verbal (scenery, gestures) of a film (Marra, 2022; cf. Remael & Reviers, 2019; Baker & Saldanha, 2009). In a similar vein and from a multimodal perspective, if AD is the result of the interaction of different semiotic systems – sound, dialogue and description – that communicate complex meanings, it should therefore be coherent for audiences and rendered in concise and vivid language so as to enable the mental reconstruction of the multimodal text (Romero-Muñoz, 2025; Mazur, 2022).

Most importantly, AD is subject to spatial-temporal restrictions: all descriptions must fit naturally within pauses and maintain synchronicity or near-synchronicity with visual content. However, maintaining synchronicity can be challenging or impossible in case of extremely fast dialogue, overlapping sound effects, scenes with non-stop visual action or elaborate scenes, e.g., a crowded battlefield, or even live and unscripted events. An interesting example could be one of the scenes from Luc Besson's *The Messenger: The Story of Joan of Arc* (1999), where Joan leads the assault on *Les Tourelles* against the English as the two armies clash in battle (Fig. 5).



Figure 5: The battle for *Les Tourelles* from “*The Messenger: The story of Joan of Arc*” (Luc Besson, 1999).

Here the visuals are fast-paced, with continuous action and a multitude of moving figures, such as soldiers advancing and arrows flying. In this scene, it would be difficult for

a describer to insert a description like “*Joan charges forward, raising her standard above the mêlée*” without interrupting or overlapping important dialogue or sound effects, because multiple visual events happen simultaneously and continuously with little natural pause. This suggests that describers are required to make careful selections and give priority to certain narrative elements while omitting others. Such selectivity places particular importance on lexical precision, as the audience relies on it to construct meaning (Vercauteren, 2016).

2.4 An overview of existing guidelines

Over the years, the need for lexical precision and careful content selection has led practitioners to seek more structured approaches, in an attempt to move beyond a fragmented landscape and towards widely recognized guidelines. In response to this need and the lack of formal training in the twentieth century, pioneers in the UK and Germany developed what are now considered seminal documents, most notably the Independent Television Commission (ITC)’s *Guidance on Standards for Audio Description* (2000) and *Dosch & Benecke Handbuch* (2004), still widely cited in contemporary research. As other European countries – including Spain, France and Poland – introduced their own services, they largely modelled their protocols on the British and German standards.

Currently, in Europe, there are eleven sets of institutional recommendations governing AD practices, including the regional frameworks of Flanders and Catalonia. These guidelines seem to function within a fragmented normative landscape though, ranging from mandated national standards in the UK and Spain, to non-regulatory practitioner manuals, e.g., in Italy and Germany. However, there is a shared understanding across the standards regarding how AD is defined and, more importantly, the four core tenets of scriptwriting, namely **what** to describe, **how** to describe it, **how much** detail is required and **when** it should be inserted (see Vercauteren, 2016). Taken together, these pillars provide a structured approach to AD and should always be in the mind of the scriptwriter to deliver well-balanced descriptions. Additionally, other important factors may also have an impact on the quality of descriptions, such as vocal delivery and voice skills (Hermosa-Ramírez, 2022; cf. Snyder, 2014; Fryer, 2019). Although there is no strict requirement for the scriptwriter to also serve as the narrator, professional delivery without

heavy regional accents is paramount and so is intonation, rhythm and synchronicity, regardless of who reads out the script. For example, the guidelines in the *American Council of the Blind's* (ACB) *Audio Description Project* (ADP) (2009) place special emphasis on vocal prosody and vocal delivery to convey the meaning of an utterance. To support this idea, the sentence "*Woman without her man is a savage*" is illustrated as a sample where, depending entirely on pausing and intonation, the meaning shifts radically between suggesting that a woman is a "savage" without a man, or that a man is a "savage" without a woman.

Returning to AD standards, comparative research shows that the existing guidelines reveal a largely prescriptive orientation functioning mainly as a collection of suggestions for discouraging poor practice (Uzzo, 2024) not adhering to a uniform structure (Bittner, 2012; cf. Vercauteren, 2007; McCall *et al.*, 2014). However, there seems to be global consensus on what descriptions should be like. Perego (2020: 119) draws on terminology previously adopted by other scholars to define AD: "meticulous" (cf. "detailed" in Remael & Vercauteren, 2011, and "accurate" in Ofcom, 2010); "visually intense" (cf. "vivid" in Remael & Vercauteren, 2011, and Snyder, 2007); "usable" (cf. "simple" in Ofcom, 2010); and "clear" (cf. Taylor, 2015); "concise" (cf. "succinct" in Snyder, 2007; cf. Uzzo, 2024). Most significantly, Snyder – one of the most prominent figures of AD research – compares it to a type of poetry with vivid and imaginative words. This is actually interesting, if we think that he himself formulated the well-known WYSIWYS motto ("what you see is what you say"), thus cementing the idea of factual, objective descriptions based on neutrality, with the audio describer being "part journalist, relaying the facts" (Snyder, 2014:65). Historically, many standards – such as the early ITC and Ofcom guidelines – upheld strict objectivity as a core tenet, whereas recent years have seen a paradigm shift toward functionalist strategies to preserve narrative coherence, as exemplified in the ADLAB project (2011-2014) and its subsequent evolution ADLAB PRO (2016-2019). These three-year projects on AD, notably financed by the EU under the Lifelong Learning Programme, were created with the aim of training AD specialists and designing guidelines that are "not prescriptive or regulatory, but strategic, and [...] depending on the context" (Remael *et al.*, 2015; Perego, 2017). Over the years, the guidelines have become seminal within MA because of their robust theoretical foundation based on a multidisciplinary approach, integrating insights from TS and narratology. The excerpt below is taken from the section about wording and

style and shows practical examples of alternative phrasings for gestures and facial expressions, each capturing a variety of nuanced descriptions:

An example of gestures from *Inglourious Basterds* (Tarantino, 2009)

- “He makes the characteristic Italian hand gesture” (gloss)
- “He makes the ‘What do you want’ gesture” Or: “He makes the ‘annoyance’ gesture” (name the gesture)
- “His fingertips pressed together, turned towards his face” (describe)
- “He makes the Italian ‘What do you want’ gesture: his fingertips pressed together, turned towards his face” (name and describe)

An example of facial expressions from *Inglourious Basterds* (Tarantino, 2009)

- “Shosanna’s eyes are wide open. She’s gasping for breath” (describe)
- “Shosanna is petrified” (name the emotion)
- “Shosanna’s eyes are wide open in terror” (describe and name the emotion)

(*Pictures painted in words: ADLAB Audio Description guidelines*. Remael et al., 2015).

Indeed, the ADLAB guidelines have contributed to elevating the status of AD to an intuitive process (Vercauteren, 2016) with a rich, multifaceted nature, rather than a linear sequence of factual information; on the contrary, they foster creative and strategic interventions, including specific jargon or socio-linguistic variants (Taylor et al., 2015), as long as intermodal coherence is maintained.

2.5 Audio Description in Reception Studies

As AVT research has recently flipped the focus toward a human-centered paradigm, the momentum in Reception Studies (RS) has placed the audience experience at the very heart of the discipline. This development is particularly evident in the field of AD, where the users’ preference and comprehension has become one of the focal points of scholarly debate, paving the way for further innovative approaches. At this point, the challenge of delineating and understanding the target user’s profile has surfaced as a central concern, often resulting in oversimplified definitions. First of all, there is no such thing as a single audience; starting from its semantic value, the term captures “an increasingly diverse and

complex reality [...] with differences of meaning” (McQuail, 1997; cf. Di Giovanni, 2020), representing a broad spectrum of people defined by age, education and media preferences (Di Giovanni, 2018). Consequently, this heterogeneity has led scholars to investigate AD reception in terms of immersion (Fryer & Freeman, 2012a; Fryer & Walczak, 2021; Hughes *et al.*, 2021; Wilken & Kruger, 2016) or engagement (Fresno, 2017), although Di Giovanni (2018) raises another important caveat regarding the audience variety, making a distinction between researching AD users as “we” versus “they”. On the one hand, the insider perspective (the “we” path) advocates research on participation and collaborative content creation (Romero-Fresco, 2018b); on the other, the outsider direction (the “they” path) prioritizes empirical, physiological data through eye-tracking and biometric measurements (Orero & Vilaró, 2012; Ramos, 2015), alongside traditional psychometric tools like surveys for quality quantification (Chmiel & Mazur, 2016).

Along these lines, much of the current literature on the reception of AD pays particular attention on the audience’s cognitive responses either *offline* – based on users’ comprehension and memory – or *online*, as a process of recipients’ meaning-making, emotions, mental imagery and involvement during the performance (Holsanova, 2022). Extending this approach further, Fryer and Freeman (2012a) investigated the impact of AD on audiences by measuring presence, that is the psychological sense of “being there”. This new exploratory approach has increasingly encouraged scholars to push the boundaries of traditional AD by challenging the long-held dictate of objectivity to measure not only presence, but also enjoyment (Fryer & Walczak, 2017). By incorporating narrative cues, such as camera movement and subjective descriptions of characters, their actions and scenes, AD may foster a deeper sense of presence, compelling the describer to explore a character’s internal and external life (Vercauteren, 2022). It is here that describing becomes an exercise of meaning-making, where the choice of vivid, narrative elements serves to enhance the user’s enjoyment and psychological engagement with the film.

2.6 The Creative Turn

Over the past years, AVT and MA have witnessed a burgeoning interest in the role of creative intervention. While early research on AD “mainly campaigned on behalf of neutral and unobtrusive descriptions to prioritize the user comprehension, recent studies

show how a consistent number of scholars have begun to evaluate AD as a tool to elicit emotions” (Pensabene, 2023:52). Creativity must not be understood as a decorative extra, but rather as a fundamental cognitive necessity allowing audio describers to move from information to connection. It is the only tool we have to share not just *what* we see, but *how* it feels to see it. In the mind of the listener, language acts as a set of instructions for building a mental model and, if objective descriptions serve as the bricks of that mental model, it is creative subjectivity that gives the “mortar”, namely the atmosphere and the depth (Johnson-Laird, 1983). Notably, one of the most challenging parts for describers is deciding what to include or prioritize; most of the existing guidelines have always promoted neutrality and do not seem to take second thoughts about this as an option. But what is neutrality? Or we should say: who determines whether a description is irrefutably objective? The answer to this question is not simple, nor is the definition of neutral unambiguous, as guidelines are drawn up by practitioners based on their knowledge and do not often provide answers beyond anecdotic resolutions (Vercauteren, 2016). Moreover, they often neglect the creative and transformative potential of MA, which is malleable, thus far from the authoritarian objectivity masked as an absolute truth (Romero-Fresco & Dangerfield, 2022; cf. Kleege, 2018).

However, defining creativity within MA remains a complex task, since AVT covers a vast array of multidisciplinary approaches, ranging from TS (creative subversion, text adaptation and localization) to multimodal and accessibility practices (abusive subtitling, user’s involvement in accessible filmmaking, or integrated AD)⁵. Along these lines, Romero-Fresco and Chaume (2022) best capture the key dimension of creative MA:

Creative MA refers to those practices that not only attempt to provide access for the users of a film or a play, but also seek to become an artistic contribution in their own right and to enhance user experience in a creative or imaginative way. (Romero-Fresco & Chaume, 2022: 84).

Most of the times, this imaginative intervention in AD is given by the use of more vivid and emotion-based vocabulary, including figurative language and neologisms, as well as the introduction of alternative storytelling strategies, *i.e.*, first-person narration (Fryer &

⁵ Abusive subtitling is a form of deliberate experimentation with the linguistic and visual dimensions of subtitles, involving creative manipulation of language and layout to give visibility to subtitlers (Nornes, 2004).

Walczak, 2017) that bring AD much closer to a form of artistic expression. In particular, the concept of *auteur* description seems to best epitomize this new creative approach that ultimately gives the describer the license to depart from objectivism to embrace vivid and emotional language (Szarkowska, 2013:383). Such descriptions, tailored for *auteur* films, incorporate the director's vision or even fragments of the published script that are not "necessarily visible on the screen, but are important from the point of view of the director" (*ibidem*), namely the ultimate creative force. Interestingly, these interpretive descriptions seem to elicit stronger emotional responses and are likely more adequate for conveying elements that are not indispensable to comprehension, but offer a better film experience (Ramos Caro, 2016; Bardini, 2020), similar to that of sighted users.

Romero-Fresco and Dangerfield (2022) push the creative approach further, advocating for the participation of blind audiences at the production table to make their voices heard as creative decision-makers rather than passive recipients. Their empowering vision brings us back to a universalist approach to AD, as a reminder of its mutable nature and interdisciplinarity. Indeed, it is this flexibility that has blurred the lines of sensory barriers, making it serviceable for sighted persons with different needs. These include elderly with cognitive deterioration, viewers susceptible to visual overstimulation, or anyone wishing to listen to a film, rather than watch it while doing something else (Morisset & Gonant, 2008).

At this point, the assumption about it being a multimodal facility for VIPs appears reductive. What began as an accessibility tool has now achieved a crossover into other environments as an entertaining service and – more importantly – as a supplementary tool for educational spaces. This pedagogical evolution lies at the core of the present study and, more generally, of current research on DAT, one of the most prolific areas of inquiry within Language Education (LE).

2.7 Audio Description in Foreign Language Education

Although AVT has long been a staple of FLE pedagogy, it has gained momentum as a supplementary pedagogical resource over the past decade, along with other established school programs. This is the result of a new elevated vision of the learner suggested by the CEFR/CV (Common European Framework Reference for Languages, 2001, 2018, 2020),

where the language learner is presented as a “social agent” acting in real life situations and carrying out multiple tasks. This Action-Oriented Approach (AoA) marks a paradigm shift, moving away from prescriptive, deficit-based evaluations and toward a proficiency-oriented perspective that prioritizes affirmative can-do descriptors over linguistic shortfalls.

A 2018 revision of the framework updated its descriptors to place greater emphasis on interaction and incorporating mediation as an additional mode of language use. As a result, the traditional notion of integrated skills – reading, writing, listening and speaking – has been reformulated within a broader model of communicative abilities, now known as reception, production, interaction and, indeed, mediation. Accordingly, these significant modifications made a pivotal contribution toward the emergence of participatory AVT as a discipline of its own, known as Didactic Audiovisual Translation (DAT, as previously introduced) and used for specific purposes. For example, dubbing can be effectively integrated into lesson plans to promote oral language production as well as peer-to-peer collaboration (Duff, 1989; cf. Talaván *et al.*, 2024). All communicative skills are equally important, but emphasis has tilted toward production skills as a necessary step to become proficient in a language. Regarding this, Incalcaterra Mc Loughlin *et al.* (2018) argue that decoding the intersemiotic connections that create meaning requires a number of digital literacies and cognitive strategies, but these are only receptive skills, if not passive. For this reason, the AoA harnesses the full potential of DAT to turn language learning into something more than simply reading subtitles or listening to AD.

In general, DAT modes can be either interlingual (from L2 into L1) or reverse (from L1 into L2); intralingual (from L2 into L2); or intersemiotic (non-verbal to L2), also used in a creative manner (*ibidem*). Among these modes, subtitling has attracted the greatest attention. Originally, it was used as a supportive activity for memory retention during language lessons, as learners could cope with the cognitive demands of processing multi-level information while watching a film in L2. Nevertheless, more concrete outcomes did not emerge until they were actively involved in the production of interlingual subtitling to improve listening comprehension (Talaván, 2010), intralingual subtitling for vocabulary development (Talaván *et al.*, 2016, 2024; Lertola & Mariotti, 2017) and reverse subtitling for writing production skills (Talaván & Rodríguez-Arancón, 2014b). The integration of these two modes has culminated in the development of several European-funded initiatives, including *LeVis* (Learning via Subtitling), *ClipFlair* (Foreign Language Learning through

Interactive Revoicing and Captioning of Clips), and the more recent TRADILEX project (Audiovisual Translation as a Didactic Resource in Foreign Language Education).⁶

AD has only recently become a prominent mode within FLE for its communication-oriented and process-based activity that make it a vibrant area for implementing innovative classroom tasks (Bausells-Espín, 2022). Early approaches were limited to passive AD, where learners used existing scripts to brainstorm verbs or other grammatical features (see Navarrete & Bolaños García-Escribano, 2022: 155), but the AoA has paved the way to new applications. As such, Didactic AD (DAD) is being used today in language classrooms to enhance vocabulary (Calduch & Talaván, 2018; Ibáñez Moreno & Vermeulen, 2013, 2017b), refine speaking skills (Ibáñez Moreno & Vermeulen, 2015a, 2016a; Navarrete, 2018; Talaván & Lertola, 2016), develop media literacy (Herrero & Escobar, 2018; cf. Bausells-Espín, 2022) and improve writing skills (Talaván *et al.*, 2024). As for the latter, intralingual script-writing activities have great potential in fostering learners' ability to understand linguistic and syntactic constructions. This is corroborated by the emergence of dedicated long-term projects including script-writing activities and mixed lesson plans, such as the ARDELE Project (*Audiodescripción como Recurso Didáctico en la Enseñanza del Español como Lengua Extranjera*) stemming from the collaboration of Ana Ibáñez Moreno (UNED) and Anna Vermeulen (Ghent University) (cf. Ibáñez Moreno & Vermeulen, 2013). Talaván *et al.* (2024) also include Didactic Free Commentary (DFC) among DAT activities as a more substantive and interventionist form of mediation. In their thematic publication, they provide an extensive account of DAD and DCF, placing emphasis on the potential benefits on language learners. An adaptation of their AD enhancement diagram is illustrated in the following page (Fig. 6):

⁶ *LeVis* is a tool developed by the Laboratory of Educational Material of the Hellenic Open University in Patras (Greece), based on the simulation of professional film subtitling (Sokoli, 2006). *ClipFlair* was a European project (2011–2014) funded by the Lifelong Learning Programme (Comenius/ICT and Languages) and managed by the Education, Audiovisual and Culture Executive Agency (EACEA) to develop a platform for Foreign Language Learning (FLL) through revoicing (dubbing/AD) and captioning. The consortium involved researchers from ten institutions: Universitat Pompeu Fabra, Universitat Autònoma de Barcelona, and Universidad de Deusto (Spain); Computer Technology Institute (Greece); Imperial College London (UK); Universitatea Babeş-Bolyai (Romania); Tallinn University (Estonia); University of Warsaw (Poland); Universidade do Algarve (Portugal); and the National University of Ireland, Galway (Ireland). *TRADILEX* is a Spanish research and development project coordinated by UNED (*Universidad Nacional de Educación a Distancia* in Madrid) and supported by the Spanish Ministry of Science and Innovation (Talaván & Lertola, 2022: 25).

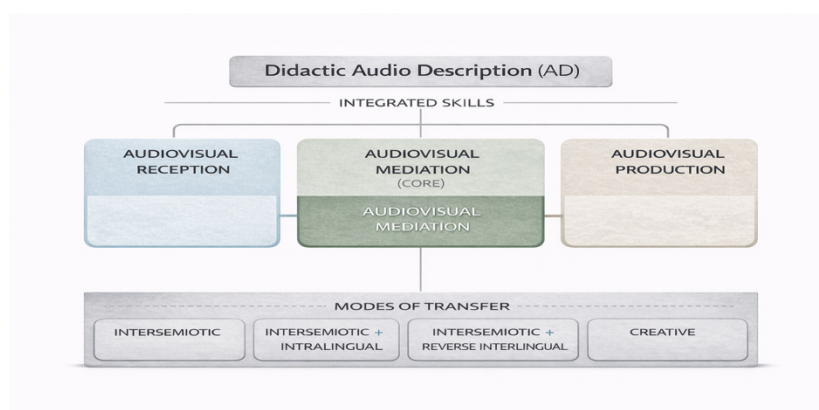


Figure 6: Didactic AD diagram adapted from Talaván *et al.* (2024).

For each mode of transfer, they suggest a number of scaffolded didactic prompts in the form of English lesson plans adapted to various levels. The samples all have a twofold purpose: write AD scripts and work on a communicative function tailored to the specific learning goal the teacher intends to cover. For example, in the Creative intersemiotic + intralingual DAD, learners practice objective versus subjective information describing things differently from what is shown. The teacher may also consider post-task activities for extra writing practice, such as watching a clip on the same topic and writing a guided short composition. Another option might be asking learners to read and record their utterances as a wrap-up moment, which is generally perceived as highly amusing. In fact, beyond the clear pedagogical aim, the participatory nature of these tasks can bolster self-motivation and maintain a positive attitude toward learning. Indeed, a proactive learning environment is the *sine qua non* for meaningful didactic progress; such progress is facilitated when learners engage in tasks that are intellectually stimulating, and intralingual DAD is certainly one of these.

At the forefront of the current landscape is the context of rapidly expanding generative AI tools, where current pedagogy must reconcile the demands of automation with creative expression. Regarding this, Professor Micol Beseghi from the University of Parma (2025) combined DAD with AI-based image generation to observe how verbal descriptions are interpreted visually by generative AI tools. Interestingly, the preliminary findings suggested that the integration of AI in AD-based tasks not only strengthens learners' motivation and metalinguistic awareness, but it also has positive effects on their critical understanding of meaning negotiation between human creativity and machine interpretation (Beseghi, 2026). Although the integration of AI tools falls outside the primary scope of this work, there is no question that these new exploratory uses of DAD warrant

further investigation within the field. However, the real takeaway here is that, regardless of the context, learners seem well aware of the pedagogical utility of such tasks, and they are equally mindful of their difficulties and what it takes to tackle them (Bausells-Espín, 2022, 2024). To do so, though, they must work in a collaborative environment with guidance and support from their teachers to settle the difference between their skills and the actual achievement, what Vygotsky (1978) defines as the Zone of Proximal Development (ZPD) (cf. Talaván *et al.*, 2024). Therefore, DAD tasks – and all DAT modes in general – can be successfully implemented if teachers receive appropriate training to maximize the existing educational resources in FLE (Pensabene, 2023).

Finally, we must not forget the importance of raising awareness on what AD means for MA, without losing sight of its pedagogical purposes. One useful way to do this is to expose learners to non-visual reception of a film or clip and ask them to make speculations about the plot in their L2. If some of them might posit correct assumptions, others may fall wide of the mark (Rodrigues Barbosa, 2013; cf. Lertola, 2019a), but in both cases, they will understand the importance of images for narrative clarity and appreciate firsthand the meaning of AD.

Chapter 3 – Didactic Audio Description in Foreign Language Education: A Practical Test

This chapter outlines the theoretical and methodological framework underpinning the present study, investigating DAD as a creative pedagogical tool for fostering advanced English-language writing and encouraging B2/C1 learners to make increasingly sophisticated lexical and stylistic choices. It first presents the pedagogical design and implementation of an experimental DAD task from the teacher's perspective before laying the methodological foundations for the subsequent analysis of the learners' scripts from a research perspective.

The discussion opens with the rationale, objectives, and implementation of the experimental task. To contextualize the activity, it introduces the three audiovisual products selected for the study, examining their narrative content, thematic concerns, and communicative purposes. The audiovisual texts are subsequently analyzed from a TS perspective through the functionalist frameworks of Reiss (1971) and Nord (1997), as adapted by Mazur (2020), in order to determine the textual characteristics of each genre. This analysis serves two complementary purposes: first, to identify the intersemiotic challenges that the source texts (STs) present to the audio describer; and second, to determine the translation method best suited to the pedagogical objectives of the experiment.

Building on this framework, the chapter describes the didactic structure of the experimental task, drawing on Talaván *et al.*'s (2024) pedagogical model. The participants, all B2-level learners of English, completed an AD script-writing task designed to promote strategic creative agency by encouraging the use of expressive adjectives, adverbs, verbs, and figurative language. The aim was to guide learners beyond literal description towards more interpretive, nuanced, and stylistically varied forms of expression.

Lastly, the final section offers an explanation of *Frazier® - VideoToVoice*, the cloud-based professional AD software used as a technological tool for the participants (the learners) during the experimental test.

3.1 Project Design: Context and Aims

Building on the historical development of DAT and DAD in FLE discussed previously, this chapter, instead, provides a full methodological account of a practical workshop with built-in tests on creative AD that was conducted with two groups of Italian learners of English.

The workshop was conducted during the training initiatives promoted under Italy's *National Recovery and Resilience Plan* (NRRP) to allocate resources and projects for the country's economic recovery following the COVID-19 pandemic. As part of a broader EU recovery plan, the proposal is mainly focused on six Missions, each addressing a key strategic area for long-term development, such as sustainability, digitalization and social inclusion. In particular, *Mission Four* revolves around Education and Research, aiming to reinforce learning systems and research pathways. It is within this context that universities play a pivotal role in achieving national goals related to education, research and innovation (Ministry of University and Research, MUR, online). To translate these institutional objectives into educational practice, the Italian Ministry of Education and Merit (*Ministero dell'Istruzione e del Merito*, MIM) has identified the *Pathways for Transversal Skills and Orientation* (known by its Italian acronym PCTO – *Percorsi per le Competenze Trasversali e per l'Orientamento*) as a strategic framework (MIM, 2019). This program, designed for upper secondary learners, aims to provide practical experience through internships and workshops in order to develop transversal skills for their future academic and professional careers.

Specifically, the PCTO workshops, held at the Department of Humanities of the University of Palermo in 2024, provided the ideal pedagogical setting to carry out the practical phase of my study. As previously indicated, the activity was included in a two-day English Language and AVT training with two groups of Italian upper-secondary school learners who engaged in creative DAD script-writing tasks to practise their L2 writing skills (English) through active involvement.

This choice echoes the recent updates to the CEFR/CV (2020) descriptors, which now formally include mediation as a mode of communication in FLE – along with reception, production and interaction – with a special focus on accessibility. Such modifications have accelerated the shift towards an AoA approach (cf. Talaván, 2020; Navarrete, 2022;

Bausells-Espín, 2022) and promoted the idea that all DAT modes, including SDH and DAD can be included within the Framework (Navarrete, 2024) for educational purposes.

More generally, DAT fosters positive attitudes toward language learning by engaging learners in constant mediation which involves reformulating, translating, paraphrasing and recreating content, not to mention that the mediation process can additionally support the enhancement of intercultural competences (Talaván *et al.*, 2024). In this study, mediation occurs both at intersemiotic and intermodal levels, where the creation of an AD script from L2 (English) to L2 (English) “requires combining both written lexical and grammatical skills to ensure accuracy and idiomaticity” (Pensabene, 2023). More specifically, the writing process here revolves around task-based creative vocabulary and subjectivity to experiment with the language on lexical, stylistic and cultural levels. The reason why creative writing was prioritized over other activities is that sometimes language learners seem to underestimate the importance of being able to put together accurate, coherent and well-structured texts, in which purposeful and well-considered lexical choices play a central role. Although Snyder (2026) argues that “99% of AD is written to be heard”, the focus here is on the production process for didactic and non-professional uses, supporting the idea that a mastery of writing competences is essential. Indeed, writing provides learners with the necessary cognitive space to evaluate and refine their lexical choices while helping them to consolidate, if not extend, their proficiency level. Taking into account B2-C1 level criteria for creative writing, learners are expected to “give clear, detailed descriptions of real or imaginary events and experiences” [...] as well as other “subjects related to their field of interest” (CEFR/CV, 2020: 67). For this reason, the learners were requested to create AD scripts in English for three different types of audiovisual products, going from passive consumers to proactive creators, namely prosumers⁷, mediating verbal and non-verbal content (Talaván *et al.*, 2024). Finally, the development of writing skills also aligns closely with the broader goals of Digital Multimodal Composing (DMC), meaning that audiovisual texts are produced by means of innovative and technological devices (*ibidem*; cf. Li, 2021) which make the learners’ experience as authentic as possible.

However, unlike other long-term initiatives, the present work was developed as an intensive, short-term study. This reduced timeframe necessitated a purposeful pedagogical approach, ensuring that the integration of AD tasks remained focused and effective.

⁷ The term “prosumer” – a *portmanteau* of “producer” and “consumer” coined by US sociologist Alvin Toffler (1980) – defines a modern paradigm shift in which the end-user goes from passive consumption to active participation in the creation and value-added processes of a good or service.

Indeed, the implementation of a creativity-driven task provided a preliminary framework for future research, exploring the impact of creative DAT-oriented strategies on vocabulary acquisition over extended periods.

3.2 Audiovisual Genres: Three Different Proposals

Following a Task-Based Language Teaching (TBLT) model, it was decided that the learners would do three script-writing exercises reflecting real-life scenarios in which they had to work in a practical situation and use their skills to meet a specific objective (Talaván *et al.*, 2024:60).

DAD, and in general AD, functions as an intralingual, AVT-based translation exercise in the form of a creative-descriptive narrative (Herrero & Escobar 2018, cf. Greening & Rolph, 2007) stimulating the learners' creativity throughout the verbalization of visual content, e.g., a film or other audiovisual product.

However, using a film, a TV program or any other type of audiovisual content as mere texts involves significant analytical challenges, if we consider that language interacts with a complex network of audiovisual elements that produce meanings (Androutsopoulos, 2012) beyond language itself. In fact, the cinematic discourse – intended as a combination of *mise-en-scène*, cinematography and editing – is a form of narration that can elicit emotional and psychological responses from viewers through visual techniques, temporal rhythm and spatial depth (see Spottiswoode, 1950). For this reason, one of the main difficulties was to identify the appropriate audiovisual material ensuring methodological coherence for the purpose of this study, even though, it was evident from the beginning that the selection would include three distinct genres, each featuring different audiovisual signifiers and their associated meanings: a drama, a talk show and a niche independent short social issue documentary⁸. This variation was intentional to have learners deal with three distinct types of AD delivery: a drama would challenge them with the description of emotional dimensions and psychological traits; conversely, a talk show would offer high-

⁸ A social issue documentary is a film that explores [...] societal problems and seeks to inform, engage, or provoke change among its audience. These documentaries often focus on topics like poverty, discrimination, environmental issues, and human rights, using real-life stories [...] to highlight injustices and encourage viewers to take action or alter their perceptions (Fiveable, Television Studies, online.)

density delivery due to frequent interactions, whereas an observational documentary introduced non-fiction focused on contemplative stasis.

After careful consideration, the choice ultimately fell on: 1) Stephen Daldry's "working-class coming-of-age tale" (Sterrit, 2000, online) *Billy Elliot* (henceforth also *BE*); 2) *The Ellen DeGeneres Show* (henceforth also *ES*), the popular US talk show hosted by comedian Ellen DeGeneres; 3) *Aagaswadi (Village in the Sky)* – henceforth also *VITS* – an observational documentary produced by the Film & Television Institute of India (FTII).

This choice sought to promote vocabulary acquisition through three different genre-specific features: emotions and figurative language (*BE*), action and physical description (*ES*), a combination of the two previously mentioned plus free narration (*VITS*).

As noted by Bondebjerg (2015), film genres are not only shaped by stylistic conventions, but also by how they tell stories and influence audience interpretation, which make these features – emotion, action and free narration – the key factors in establishing AD strategies. Based on this, the audiovisual content included three short clips in which different levels of interaction alternated to stimulate the learners' creativity. Therefore, while the *BE* clip was marked by emotional density with zero interaction, in the *ES* fragment dialogues complemented a vivacious atmosphere, before concluding with *VITS*, in which visual stillness deliberately contributed to a graver, nearly elegiac tone. These contrasting features were specifically selected to elicit three different approaches to AD and challenge the participants to adapt their descriptive abilities accordingly.

With reference to cinematic work, over the years popular films have increasingly evolved into complex narrative structures (Cutting, 2019b), whose characters' intentions and emotions must be interpreted through gestures and body language; this is the case for the *BE* clip, which mainly relied on non-verbal cues and emotional performance. Conversely, *ES* combined the non-verbal cues with frequent locutionary and illocutionary acts (Austin, 1970) resulting from humor and irony, two key elements found in satire and comedy. Lastly, *VITS* comes across as a non-fiction ethnographic documentary following an observational style with "measured pace [...] and laconic narratives" (Marano, 2022:39). Here, the raw-footage technique allows for the audience to be propelled to a circular dimension of otherness, resulting in a counter-narrative through which marginalized content – such as people coping with extreme environmental conditions – can be disseminated to bring attention to certain minoritized communities.



Figure 7: Posters and factsheets of the three selected audiovisual genres.

On this basis, the following sections establish a TS- informed approach to text types within the audiovisual domain. Subsequently, they provide a general overview of the three audiovisual products, their narrative themes, followed by analyses based on a functionalist approach.

3.3 A functionalist approach to text types

Before establishing the thematic and narrative landscape of the three clips, I will now introduce a categorization of the audiovisual products inspired by Reiss's (1970) and Nord's (1997) functionalist models, later adapted by Mazur (2020). This approach offers a useful framework to identify effective methods for creating a translated text, as AD itself is a type of translation (Mazur, 2020). The distinction here is that the transposition occurs at an intersemiotic level, not moving between two linguistic systems, but from the audiovisual product – the ST – to a verbal rendition of the AD script, namely the target text (TT). Following this rationale, the three genres are deconstructed according to functionalist parameters to delineate the features of each text type and their function (*Skopos*). Although these theoretical labels were not disclosed to the participants, they make it possible to explore the intersemiotic challenges of the STs and to determine the translation strategies most suitable for didactic purposes.

Before moving forward, some theoretical clarification is required. It is necessary to point out that the functionalist turn adopted by Reiss was originally intended to ensure a more systematic assessment of translation. Reiss sees the text as the level at which communication happens and equivalence must be secured (Munday, 2016:115; cf. Reiss 1977/1989) starting from three functions of language: 1) informative, 2) expressive and 3) appellative. These functions each relate to distinct language “dimensions” and to the text types in which they are typically used.

Reiss (1989) suggests “a classification of text types into informative (content-focused, e.g., a reference work), expressive (form-focused, e.g., a poem), operative (appellative-focused, e.g., an advertisement)” (Mazur, 2020:227), according to their language functions. However, not all texts belong exclusively to one type or genre and may be hybrid in nature, fulfilling more than one communicative function. A biography, for instance, serves both informative and expressive functions, whereas a religious speech merges expressive rhetoric with an operative purpose (see Munday, 2016) – a dual nature that is equally baked into the audio-medial text, also known as multimodal text. This brings us back to the question of AD, which Mazur (2020) conceptualizes as a continuum of multimodal texts also including the narrative (e.g., films, biopics) and entertaining (e.g., game and talk shows) typologies. Here, the hybrid character of such texts is evident once more, seen in their blurred boundaries and overlapping functions. Put differently, these types are not “clear-cut” categories but indicate dominant communicative tendencies, as audiovisual texts are usually hybrid, with one function standing out. For this reason, films may exhibit expressive qualities as much as art films may still retain some narrative elements; similarly, commercials and reality shows may incorporate storytelling or stylistic features (*ibidem*).

Therefore, taking into account Reiss’s classification of text types and Mazur’s multimodal framework, I propose a re-elaboration of the model (Table 1):

Category	Reiss			Mazur	
	Informative	Expressive	Operative/Persuasive	Narrative	Entertaining
Language function	Representation of facts and information	Expression of the sender’s attitude	Appeal directed at the receiver	Storytelling	Enjoyment
Language dimension	Logical	Aesthetic	Dialogic	Narrative /sequential	Engagement-oriented
Text focus	Content-oriented	Form-oriented	Effect-oriented	Plot and character-oriented	Audience engagement

Category	Reiss			Mazur	
	Informative	Expressive	Operative/Persuasive	Narrative	Entertaining
TT goal	Transmit referential content	Preserve artistic form	Elicit desired response	Coherent narrative development	Retain pacing and engagement

Table 1: Author's integration of functional characteristics of text types according to Reiss (1971/2000) and Mazur (2020).

Beyond defining text types, the three audiovisual products will be analyzed following the most salient non-linguistic extratextual and intratextual determinants, based on Nord's (1997) and Reiss's (1971/2000) models, with an integration of Mazur's (2020) multimodal AD framework. This will help shed light on two aspects for (audiovisual) translation training, namely the importance of what Nord (1997) calls the "Translation Commission" and the role of the ST analysis. The extratextual features will include:

- intended text functions;
- audience;
- medium;
- addressees (sender and receiver);
- motive (why the film is being described);

Instead, the intratextual aspects will break down the ST at the macro and microlevel and will include the following elements:

macrolevel

- genre;
- text composition (thematic structure, logical relations);
- subject matter (how culture bound it is to the source and target audience);
- protagonists

microlevel

- specific visual elements;
- filmic cues;
- affective implications.

Once the text types and their functions are established, we will discuss some possible translation methods – here understood as approaches to AD. At this point, it should be remembered that the audio-medial text type mainly relies on a supplementary method. This implies that when one of the key semiotic modalities is partially received, any translation (be it linguistic or intersemiotic) must compensate for these deficiencies, which

is when AVT mainly resorts to multimodality (see Taylor, 2020:91); in the case of AD, this process involves the verbal rendition of visual elements that would otherwise be inaccessible for BVI. Indeed, the audio-medial text appears to be offering an overarching framework that cuts across all text-based methods, since AD in itself is a compensatory strategy in this sense, regardless of the ST type. Furthermore, as AD is multimodal and intersemiotic, it is more accurate to distinguish between the AVST and the AVTT, thereby ensuring that description strategies are tailored directly to the audiovisual source. For example, informative AVST should be described in clear, straightforward language, avoiding subjective interpretations and making salient information explicit. Expressive AVST should aim to preserve aesthetic qualities by highlighting visually and emotionally meaningful cues (such as facial expressions, lighting, color, and pacing), and, where appropriate, by using evocative or nuanced wording. Operative AVST, in turn, should be described in a way that enhances the impact of attention-grabbing elements (such as product features or physical attractiveness).

In narrative AVST, describers should follow a content-led approach to maintain coherence, ensuring that characters, settings, and actions remain easy to recognize throughout the sequences. Finally, for entertaining content, describers should preserve rhythm and keep the audience engaged; this may involve paying close attention to pacing and action, as many entertainment programs tend to run at a brisk pace (see *ES*). The table below provides an exemplification of the most common AVST types and their general implications for AD strategies:

Text Type / Genre	Primary Function	Secondary Functions	Typical AD Macro Strategy
News broadcast	Informative	Operative, expressive	Clarity, precision, prioritization of key facts
Documentary	Informative	Expressive, narrative	Content preservation + selective aesthetic detail
Feature film (drama)	Narrative	Expressive, informative	Narrative coherence + mood-sensitive description
Romantic film	Expressive	Narrative	Emphasis on affect, character interaction, atmosphere
Comedy series	Entertaining	Narrative, expressive	Timing, support for humour
Advertisement	Operative	Expressive, informative	Selective emphasis, attention-guiding, impact
Sports broadcast	Informative	Entertaining	Fast-paced, action-focused, prioritisation of events
Children's program	Entertaining	Informative	Simplicity, clarity, engagement

Text Type / Genre	Primary Function	Secondary Functions	Typical AD Macro Strategy
Theatre performance (AD)	Expressive	Narrative	Atmosphere, staging, character dynamics
Museum video/exhibition	Informative	Expressive	Clear explanation + salient visual detail

Table 2: Functional classification of audiovisual text types and implications for macro-level audio AD strategies (author's elaboration).

That said, most texts (be they multimodal or not) are hybrid, meaning that they may also serve more than one *Skopos*, even though the transmission of the primary purpose is the determining factor by which the TT is judged (Reiss, 1977/1989:109; cf. Munday, 2016, Mazur, 2020). These functions are also context-dependent and may vary depending on the audience and their expectations. Therefore, “the overriding *skopos* for the AD will be to transfer those ST constituents that will render the AD as a cohesive and coherent narrative whole, at the same time respecting the intersemiotic correspondence between the film and the AD” (Mazur, 2024:519). Needless to say, the translator should prioritize the leading function without losing sight of the secondary ones.

Returning to the film *The Messenger: The Story of Joan of Arc* as an example, we may say that its primary function is narrative, as the story follows Joan's legendary journey during the final phase of the Hundred Years' War. At the same time, the film takes on a strong expressive dimension under Luc Besson's visionary direction, which calls for a content-oriented macro-strategy in AD that takes into account the plot and characters, as well as the expressive elements. In other words, the description should still do justice to actions, character identification, orientation in battle scenes, while also including affective implications (e.g., Joan's expressions of fear, determination or spiritual turmoil), filmic cues (e.g., rapid editing, lighting contrasts, close-ups), not to mention specific visual elements (e.g., late medieval costumes, settings, and physical interactions) that contribute to the emotional and dramatic impact of the film.

In the context of this study, AD moves learners toward a more critical understanding of the AVST teaching them to identify its most salient semiotic features, before they can implement descriptive strategies in line with the *Skopos* of each AVST. At its most effective, such a *Skopos* is served when the description is not only coherent, but also tuned to the AVST main communicative purpose, so the didactic focus on AD does not exempt learners from this requirement. This study adopts a more flexible approach that validates creativity

and subjectivity not as deviations from the *Skopos*, but as purposeful and functionally motivated strategies, as long as they remain coherent with the AVST and do not obscure visual information.

However, the functionalist framework was not presented to them as such; still, there is an implicit echo of a function-oriented model in the entire process that revolves around a purposeful and selective use of the audiovisual clips. In fact, the script-writing activity pushes learners to be creative, while also improving their lexical and intercultural competences, provided that the majority of them owned a B2 level of English. More precisely, it is a scaffolded creative process that encourages them to draw on a wide range of grammatical resources (verbs, adjectives and adverbs), along with stylistic flexibility to bring out the functional purpose of the AVST.

3.4 *Billy Elliot*

As anticipated in Section 3.2, *BE* is Stephen Daldry's coming-of-age comedy-drama, starring British actors Jamie Bell as Billy and Gary Lewis, in the role of Jackie Elliot (Billy's father). The film is "set against the brutal backdrop of the 1984 miners' strike [...]" and tells "the story of a working-class lad determined to overcome prejudice and poverty to pursue his dream of becoming a professional dancer" (Sagar, 2025, BBC, online).

Critically, the film garnered numerous accolades, including three *BAFTA Awards*, with Jamie Bell becoming the youngest winner of *Best Actor in a Leading Role*. This remarkable success also extended beyond the screen, resulting in a theatrical adaptation in 2005 with music production credited to the acclaimed British singer-songwriter, Sir Elton John.

Synopsis

1984, Everington, County Durham, North of England. A year after his wife passed away, Jackie Elliot is still lingering in sorrow. As a coal miner coping with budgetary pressures, he joins his co-workers on strike to protest issues with the union. With Tony, his eldest son acting as a key union enforcer, the striking workers are reluctant to return to work, fearing potential retaliation. Meanwhile, Jackie believes that sending his youngest

son Billy to boxing classes will help him to build strong discipline and boost his confidence, but Billy's life is changed forever one day, when he accidentally stumbles upon a ballet lesson and is instantly captivated. Without hesitation, he secretly starts attending ballet dance classes with Mrs Sandra Wilkinson, a disheartened teacher, whose motivation is soon refreshed by the boy's hidden talent. When Jackie is confronted with the truth, Billy's passion is challenged by his father's cemented idea that ballet dance is predominantly a feminine expression. Despite Jackie's aversion, Mrs Wilkinson is highly supportive of Billy and offers him evening private lessons to show up for an open audition at the Royal Ballet School of London. However, as tension rises among the striking coal miners, Tony is arrested after a heated quarrel flares up, which inevitably causes Billy to miss his audition.

When Mrs Wilkinson confronts Billy's family about the painful experience, it comes as a shock to Jackie that his son has not quit dancing. It is during the Christmas holidays that Jackie becomes aware of his son's talent, when he finds him dancing at the gym with Michael, a homosexual guy with whom Billy has always shared a strong connection, taking his distance from his father's narrow-mindedness. Now determined to support his son, Jackie does everything he can to raise money for an audition in London, despite Mrs Wilkinson offering to help him.

At the audition, Billy's unrefined technique does not fail to impress the judging panel, to whom Billy becomes candid about his passion, comparing himself to electricity.

According to Daldry himself, "the scene is an interesting exploration of French sociologist Pierre Bourdieu's theory of cultural capital, which suggests that someone's [...] mannerism and style of speech [...] demonstrate their social position in society" (cf. Mohdin, 2022, online):

Though the speech is raw and not as articulate as he wants, Billy's response is very honest and very emotional; the scene is 'a struggle of struggle'. But it's not just about Billy finding his voice, it's about the dad as well. He has never heard his son express what he feels like (Stephen Daldry, 2022).

However, his escalating nervousness culminates in a violent physical outburst against another boy, leaving him feeling hopeless and discouraged about the outcome of the audition. To Billy's surprise, the Royal Ballet School sends him an acceptance letter which marks the beginning of his academic path in London to pursue his long-time dream.

A time skip of 14 years shows Billy, now aged 25, preparing for his stage performance of the *Swan Lake*, as his family and Michael take their seats in the theatre. When Billy finally makes his breathtaking entrance, emotion rushes over Jackie, bringing tears to his face⁹.

The real story of British ballet dancer Philip Mosley was the inspiration behind this emotionally-dense film and was obviously enriched with fictional elements. Contrary to Billy, who initially lacked family support, Philip never had to sneak in through the ballet gym door and, before he turned 10, he had already won a scholarship at the Royal Ballet School. Still, Philip himself claimed his life story was the catalyst for the creation of Billy's character and, while their journey to their dream was not always comparable, their final achievements definitely rank among the most inspirational success stories of all time.

3.4.1 Exploring the film's Themes and Meanings

Interestingly, the final scene of the film, showing Billy leaping into the air carries a deeper meaning than simple triumph. The film's increasing non-diegetic structure highlights the character's arc and the ultimate payoff of the story; if, on the one hand, the view of Billy jumping conveys his symbolic alignment with the *Swan Lake*'s character, on the other hand, this slow-motion *grand jeté* encapsulates the film's themes of resilience and victory over prejudice about the stigma on male dancers (Prado Martinez, 2015, online) as well as Billy's sexual orientation.

As for the former, not only can the film be considered as a motivational story for those who decide to go after their aspirations, but it has also been liberating for male ballet dancers, who were somehow pressured to cope with the effeminate stereotypes of their art form. As Haltom and Worthen (2014:761) observe, male dancers often "self-enforce heteromascularity in both their body movement and social patterns" as a way to combat social stigma (cf. Mennesson, 2009). The latter is a contentious issue, because the film explores the subject of masculinities, depicting the community eventually coming to terms with homophobia, even though Billy's sexual orientation remains deliberately ambiguous (Buckingham, 2017, online). However, while such topics deserve in-depth analysis, they will not be considered further on. Conversely, what is central for this study is the intense

⁹source: this synopsis has been reconstructed following the example of www.gradesaver.com

emotional resonance permeating the entire film, something that the participants immediately noticed. The emotive dimension plays a pivotal role and was introduced to the participants as the main creativity-enhancing element for developing AD strategies.

Generally speaking, *BE* can be defined through its multifaceted emotional depth, ranging from passion to fear and sadness, from determination to tension, not to mention hope and happiness. Among the vast array of emotionally charged sequences in the film, the acceptance letter scene significantly stands out as an emblematic moment of emotional climax, when Billy finally finds out that he has been admitted to the Royal Ballet School of London. In the scene, Billy returns home from school and finds his family waiting for him in the kitchen, nervously looking at an envelope resting on the table.



Figure 8: Billy goes through emotional turmoil.

There is a strong sense of anticipation and tension, as both the characters and the audience are unaware of the outcome. This intense mood, though, eventually culminates in relief and uncontained happiness when all the characters learn about the positive response.

However, the learners were asked to work on the first part of the scene, which was accordingly cut seconds before the letter was opened and they were not shown the remaining sequence until their AD had been completed (the final moments of the scene were part of a post-task exercise. For further clarification see Chapter 4, Section 3.11.2).

This scene was ultimately chosen to teach learners how to experiment with verbal creativity by manipulating words and exploring a variety of syntactic structures to develop their vocabulary and, possibly, their understanding of language nuances. The main

challenge was to detect emotional intensity through non-verbal cues before processing these inputs at an intersemiotic and intermodal level and translating them into verbal written outputs.

3.4.2 Text composition of *Billy Elliot*

With the parameters outlined in Section 3.3, the discussion now moves to scrutinize the textual composition of each audiovisual text, as well as their related issues and possible AD strategies. However, an important caveat is required: I will not take into account the broader commission of the AVST as a whole; instead, I will only refer to the functional characteristics of the selected excerpts to discuss AVTT strategies, which will be evaluated in Chapter 5. Starting from *BE*, the functional characteristics are set out in the following table:

AVST: Billy Elliot - The Acceptance Letter Scene		
Level	Parameter	Analysis
Extratextual	Intended text functions	Narrative / Expressive (tension building, with resolution withheld)
	Audience	BVI viewers; in this study, learners producing AD for didactic purposes
	Medium	Audiovisual film excerpt rendered through AD (in written form)
	Addressees (sender/receiver)	Sender: film + AD describer; Receiver: BVI audience (intended) / learners
	Motive	To provide access to key narrative and emotional elements (intended); pedagogically, to foster selective description and expressive language
Intratextual – Macrolevel	Genre	Feature film (drama)
	Text composition	Scene structured around anticipation; the letter functions as a narrative pivot, but its content is not revealed within the clip
	Subject matter	Family expectations, personal aspiration, social class
	Protagonists	Billy and his family (father, brother, grandmother), with emphasis on relational dynamics
Intratextual – Microlevel	Language function	Non-verbal communication: silence, gaze, minimal movement
	Language dimension	Cinematic / aesthetic (meaning conveyed through visual and filmic cues)
	Specific visual elements	The letter, body positioning, gestures and expressions
	Filmic cues	Pauses, close-ups, stillness
	Affective implications	Suspense, anticipation, emotional tension without resolution

Table 3: Functional characteristics of The Acceptance Letter Scene (author's elaboration).

The letter scene fulfils a narrative function, because it clearly shows a key development in the plot and marks a turning point in Billy's life. At the same time, it has a strong expressive dimension stemming from the characters' emotional limbo. Interestingly, the scene revolves around the letter, which functions as a strategic pivot despite remaining entirely unspoken, as well as the interaction among characters. So, meaning is conveyed through non-verbal cues, such as gaze and small gestures rather than dialogues. As we know, the clip is cut before the final reveal, leaving Billy fraught with anticipation; as a result, tension is progressively built but not resolved. Additionally, the affective tone of suspense and anticipation is corroborated by static shots, prolonged pauses and subdued lighting.

The main big hurdle in the scene is managing the narrative tension, as learners may find it difficult to resist the urge to over-articulate what remains visually implicit. At the same time, the sequence is defined by what Kurzon (2007) terms "situational silence", which may initially suggest absence, becoming increasingly apparent once it is manifested. Learners should rise to the aesthetic challenge of the sequence in a creative way, avoiding descriptions that risk flattening what is, in fact a significant moment. Additionally, the presence of culture-specific presuppositions, such as Billy's school uniform or the domestic setting are some aspects that relate to their intercultural knowledge and may sometimes escape their notice.

On a macro-level, the purpose of the description is to render the build-up of tension without prematurely disclosing the outcome. On the micro-level, silence becomes a carrier of meaning rather than a mere absence. Learners should therefore work with silence providing essential visual cues – such as the letter, the facial expressions and body language – while maintaining the scene's suspended quality. Other significant inputs may come from elements of the domestic setting, including clothing and interior details, which help delineate the socio-cultural context and should therefore be included in the description as culture-bound elements.

There are, of course, various strategies for producing different types of viable AD, as suggested by recent research on AD and Film Studies (cf. Mazur, 2024), but in the present study, I will consider a number of possible tactics for first-time describers in a didactic context, adapted to the three genres.

These features suggest a range of possible descriptive strategies, such as the spacing of descriptions to relay narrative tension (e.g., *"He stands still". The letter in his hands. Unopened.*"), which can be reinforced through fragmented syntax and delayed

information (e.g., “*He pauses. Waits.*”; “*He opens the letter. His eyes scan the page. He stops.*”), along with strategic emphasis on key elements (“*In his hands, the letter*”), so that nothing is given away early on. The use of evocative elements is not only permitted but strongly encouraged in response to the scene’s expressive tension, leaving plenty of room for creativity. Consequently, learners are prompted to use adjectives and adverbs, moving beyond neutral phrasing (e.g., “*He grips the letter tightly, hesitantly*” instead of “*He holds the letter*”). A potential pitfall is that they may draw erroneous inferences about the characters’ emotions; on the other hand, the task may be facilitated by the repetition of facial tropes, although these are frequently genre-dependent (Igareda, 2011).

A possible solution is to combine action with emotion (e.g., “*He hesitates, uneasy*”) and alternate sentences like “*A tense silence lingers between them*” with nominal phrases like “*A tense silence between them*”. Indeed, figurative language can help bring out implicit emotions, so it is preferable to suggest prompts of set phrases that can assist the creative writing process.

In line with this, the main descriptive dimensions and associated challenges of the scene can be summarized as follows:

Analytical dimension	Descriptive focus	Associated challenges
Narrative progression	Sequence of actions (reception and opening of the letter)	Premature disclosure; over-explication of visually implicit information
Narrative tension	Hesitation, fear, and hidden outcome	Difficulty sustaining suspense; anticipatory or resolving descriptions
Silence (situational)	Pauses, and absence of dialogue	Reduction to neutral formulations that fail to convey communicative significance
Focalisation and emotional cues	Glances, gestures Facial expression, characters’ emotions	Overlooking subtle cues; disclose psychological traits
Contextual elements	Domestic setting and culturally marked details (e.g., school uniform)	Omission of relevant cues or inclusion of non-functional detail
Lexical and stylistic realization	Creative use of adjectives, adverbs, figurative language	Overly neutral renderings vs. excessive expansion

Table 4: Descriptive dimensions and associated challenges in the Billy Elliot acceptance letter scene.

However, these potential solutions are by no means prescriptive, but only indicative of the challenges that learners may encounter and the strategies they may adopt; their actual implementation in the scripts will be examined in Chapter 5 based on creativity levels.

3.5 *The Ellen DeGeneres Show*

ES was a US variety and talk show hosted by comedian Ellen DeGeneres and produced by Telepictures Productions. The show, which ran for nineteen seasons – from 2003 to 2022 – must not be confused with *The Ellen Show*, a single-camera sitcom that aired in 2001, in which DeGeneres played Ellen Richmond, a fictional character facing a breakup. It was after the first season that the sitcom finally evolved into an unscripted, host-centered show, featuring Ellen DeGeneres as herself interviewing celebrities, staging pranks and sharing uplifting stories.

3.5.1 Exploring *The Ellen DeGeneres Show*’s Themes and Meanings

The show blended comedy, musical performances, celebrity interviews and inspiring personal stories altogether, also including games and prize giveaways. This concept was complemented by a variety of jokes and signature sound effects (e.g., Ellen saying “Aww snap!”), which made it one of the most popular talk shows worldwide. There were also a multitude of recurring segments often featuring audience interaction, such as *Know or Go*, in which three audience members answered questions on different topics, *Oh Hair No!* featuring fans sending their funny haircuts, *Spill the Tea*, in which Ellen tried to match secrets to five audience members and asked them to reveal them once the correct matches were made. These and other segments helped maintain the program’s brisk pace and a warm atmosphere and were often balanced with slower and heartfelt moments. Those included both celebrities and ordinary people telling their stories about overcoming adversity or manifesting their dreams, as was the case of a school senior to whom Ellen offered a \$20,000 scholarship following his objection to his school’s hair policy, or a child who was given the opportunity to sign with a record label after his rendition of one of Lady Gaga’s songs had gone viral on YouTube¹⁰.

¹⁰ In 2020, Trinidad-born student and Texan resident DeAndre Arnold faced a dispute with his school over his long dreadlocks and received a scholarship, courtesy of Ellen DeGeneres and singer Alicia Keys (Griffith, 2020, online); in 2010, 12-year-old Texas-based student Greyson Chance uploaded his cover of Lady Gaga’s *Paparazzi* on YouTube and quickly became a sensation. Reportedly, he received a phone call by the singer and later signed a recording contract (Haasch & Cohen, 2022, online).

Initially grounded on playful segments, talent showcase sections and celebrity moments, *ES* gradually integrated social themes, philanthropy and LGBTQ+ advocacy during its peak popularity (2007-2015). Unfortunately, toward the end of its run, the show faced public scrutiny related to workplace issues surfacing from internal investigations, which inevitably affected its public image leading to its cancellation in 2022.

Interestingly, its multi-purpose formula had a considerable impact on television to such an extent that DeGeneres used her platform to set a precedent for talk shows to engage with important societal themes while maintaining an uplifting tone (online).

However, much of her legacy on the entertainment culture is a result of her rhetorical strategies, mainly based on exaggeration, facetiousness and verbal humor (Tianli *et al.*, 2022). The latter is central in her monologues which often combined techniques such as incongruity, superiority and relief (Faula *et al.*, 2022). Indeed, it was not uncommon for Ellen to say unexpected lines or treat her famous guests with gentle mockery, not to mention her witty tension-relieving jokes:

I thought that you had to be a famous alumnus, alumini, aluminum, alumis; you had to graduate from this school. And I didn't go to college here, [...] I didn't go to any college at all, [...] but look at me, I'm a huge celebrity! (Haskell, 2018).

ES was developed around a modular narrative structure, with each episode opening with a monologue serving as an ice-breaker to set the mood. The central portion, usually featuring celebrity or human-interest stories alternated with games and musical performances, followed by a closing send-off, where Ellen delivered a final joke. The show frequently incorporated thematic segments around holidays or seasonal events, like Halloween or Christmas, thus creating a cohesive focus for the entire episode which inspired sketches, games, and guest appearances.

After careful examination, an excerpt from a Halloween-themed episode was deemed appropriate for the present study and eventually chosen as the second AD writing task. The episode belongs to Season Ten, which aired in 2012 and included a segment in which Ellen introduces a group of children dressed in a wide variety of creative Halloween costumes for entertainment purposes. The children walk across the stage, showing off their outfit to the live audience, while Ellen asks brief questions and reacts to their gestures.

Occasionally, she surprises the children with sweets and makes playful comments to highlight the comedic atmosphere. However, the participants were asked to work on the first portion of the segment, in which three costumes are introduced onto the stage. The first boy impersonates Clint Eastwood, followed by a girl carrying a chair and resting her face in the oval cutout of its backrest, a clever reference to Clint Eastwood's surprise speech at the 2012 Republican National Convention, during which he addressed an empty chair pretending that Barack Obama was sitting in it. Eastwood, a self-proclaimed libertarian, improvised a one-sided conversation with the former President, asking him elucidations on his allegedly unfulfilled campaign promises. The speech became one of the most controversial moments at that time and inevitably received polarized reactions with conservatives praising his creative effort and other critics branding the empty-chair act as openly disrespectful. However, the moment had a huge cultural impact and resonated beyond the convention audience. It soon became a popular parodied skit in the media, also sparking discussions about celebrity involvement in politics and showing evidence that famous figures can have a considerable influence on the cultural discourse.



Figure 9: The Clint Eastwood costume moment.

Finally, the third boy offers another humorous moment when he appears slightly disoriented in his classic magician suit with a top hat and a cane. He is encouraged to take off his jacket to reveal a padded bare-chest costume, a direct allusion to *Magic Mike*, a newly released motion picture at that time. The film was directed by Steven Soderbergh and starred Channing Tatum as Mike Lane, a stripper from Tampa pursuing his personal goals in the entertainment industry. Like Clint Eastwood, Channing Tatum increased his media prominence for his dancing skills and charisma, drawing widespread attention on US talk shows and social media.

Overall, all these caricature-like features prompting on-stage action deserve special attention in AD, if we consider that “when audiences process and interpret a story, they will look at the actions that are being performed and combine them with other information [...]” (Remael *et al.*, 2015). For these reasons, *ES* poses its own genre-specific challenges for AD, ranging from dynamic flow to visual humor, let alone the characters’ physical description.

The costumes closely reflect a couple of culture-specific references (henceforth CSRs) – be they film characters or cultural sensations – that should be either interpreted explicitly or simply described objectively. Such references are generally rendered in AD through a number of strategies that draw on the describer’s intercultural competence (Jankowska, 2022, cf. Remael & Vercauteren, 2010). Although high-school learners are not generally expected to have complete encyclopedic knowledge to detect the source CSRs (see Rizzo & Spinzi, 2023), these culturally marked elements provided some interesting cues for developing their AD.

3.5.2 Text composition of The *Ellen DeGeneres Show*

To introduce the Halloween Costumes segment from *ES*, we shift from the emotional intensity of *BE* to a completely different generic landscape: Television Variety/Talk Show. Unlike the narrative depth of the first AVST, this genre is more fast-paced and relies more on extratextual knowledge, *i.e.*, culture-bound references. Before discussing possible strategies for the AVTT, an overview of the functional characteristics is provided in the following table:

AVST: The Ellen DeGeneres Show - The Halloween Costumes Segment

Level	Parameter	Analysis
Extratextual	Intended text functions	Entertaining (primary) / Operative (secondary), centered on humor and audience engagement
	Audience	General TV audience; in this study, learners producing AD for didactic purposes
	Medium	TV talk show excerpt rendered through AD (in written form)
	Addressees (sender/receiver)	Sender: program + host (Ellen) and AD describer; Receiver: TV audience (intended), with learners acting as scriptwriters
	Motive	To entertain and engage the audience through humor and interaction; didactically, to foster creative and expressive language use

AVST: The Ellen DeGeneres Show - The Halloween Costumes Segment

Level	Parameter	Analysis
Intratextual – Macrolevel	Genre	Talk show / entertainment program
	Text composition	Sequence of humorous costume suggestions
	Subject matter	Culture-bound references (e.g., Clint Eastwood, <i>Magic Mike</i>), humor based on parody and visual gags
Intratextual – Microlevel	Protagonists	Ellen (host) and the children
	Language function	Meaning conveyed through humorous speech, timing and audience response (laughter, applause)
	Language dimension	Engagement-oriented / dialogic
	Text focus	TV dynamics; comedic effect
	Specific visual elements	Halloween costumes, visual incongruity (e.g., child vs Clint Eastwood)
	Filmic cues	Static shots
	Affective implications	Amusement, lightness, shared humor, audience involvement

Table 5: Functional characteristics of the Halloween Costumes excerpt (author's elaboration).

The primary function here is entertaining, typically associated with light game shows, contests and reality shows (Mazur, 2020), whose main purpose is to amuse and engage the audience. The segment also serves an operative function – or persuasive, in Mazur's words – aiming "to exert an influence on the viewer's behavior" (*ibidem*), or elicit a response from the audience. In a sense, the costume parade reinforces the show's brand identity, while attempting to draw laughter from both the studio audience and the viewers at home. Unlike commercials, here persuasion is directed towards triggering audience enjoyment rather than an actual call to action.

From a didactic perspective, the segment presents a number of challenges related to humor, rhythm and the interpretation of culture-bound references. On a macro-level, learners must insert the description at the exact moment the visual joke happens to ensure the final receiver experiences the irony at the same time as the sighted audience. This is particularly demanding in fast-paced moments, where visual and verbal cues follow one another. On a micro-level, a further challenge consists in the description of the two most relevant "facts", namely the Clint Eastwood - Chair moment and the Magic Mike gag; in the former – a political satire reference – the comic effect relies entirely on visual and situational elements, while in the latter, humor arises from the visual inconsistency between the child's magician costume and the adult Magic Mike stripper reference. We may assume that the humorous effect is purely semiotic, relying on the contrast between the signifier, e.g., a small child wearing extra-large clothes or a suit, and the signified, be it a celebrity politician in his eighties or a male stripper. Now, we know that humor represents one of the greatest

challenges in the various modes of (audiovisual) translation, especially when the humorous effect is created through visuals (Martínez Sierra, 2010). However, a fuller discussion on the topic would fall outside the scope of this thesis, so the description here will be limited to making this incongruity explicit. For example, instead of writing “*a boy in a suit*”, learners must highlight the specific irony triggers, such as the grey wig, the oversized chair, or the tear-away suit for the Magic Mike costume. Similarly, they should focus on the choice of vivid verbs for gestures and actions to facilitate the intermodal transfer of visual meanings to verbal ones.

Drawing on Nord’s concept of presuppositions, this case presents an interesting scenario: both the AVST and the AVTT are in the learners’ L2, yet the “translation” remains predominantly source-oriented. As a result, we cannot assume that sender and receiver share the same presuppositions, particularly regarding culture-bound references and humor. A further challenge therefore concerns the learners’ limited intercultural competence as well as their lack of familiarity with the comedic conventions required to fully appreciate these nuances. On the other hand, because the task is mainly didactic, full adaptation for a target audience is not the primary goal; instead, the focus remains on the learners describing and negotiating these culture-bound references. The main descriptive dimension of the segment and the challenges can be synthesized as follows:

Analytical dimension	Descriptive focus	Associated challenges
Narrative progression	Sequence of entrances, costume reveals and reactions	Risk of overloading or under-specify description
Pacing	Fast-paced interaction, rapid exchanges, audience responses	Omission of key elements
Humour and performativity	Comedic gestures, playful tone	Flattening of tone
Intertextual references	Cultural references (e.g., Clint Eastwood, Magic Mike)	Lack of shared cultural knowledge; uncertainty in explicating or retaining references
Costume description	Visual details of costumes and their recognisable features	Over-specification vs. insufficient detail
Focalisation	Physical appearance and visual incongruity	Capture performative cues
Lexical and stylistic realisation	Creative use of verbs, adjectives	Overly neutral renderings vs. excessive expansion

Table 6: Descriptive dimensions and associated challenges in the Halloween Costumes segment.

As for *BE*, the implementation of AD strategies will be explored in Chapter 5.

3.6 Aagaswadi – Village in the Sky

VITS is a poignant social and environmental twenty-minute-long documentary directed by India-native director Ramesh Holbole in 2018, while he was pursuing his MA in screenwriting and directing at the *Film & Television Institute of India* (FTII). In 2020, the short film received international recognition for its advocacy of social issues among niche film festivals, with screenings at the *Nara International Film Festival* (Japan), the *Lago Film Fest* and *SoleLunaDoc Film Festival* (Italy), among others. The documentary also won the *Paul Călinescu Award* for Best Documentary Film at the *22nd Cinemaiubit International Student Film Festival* in Bucharest, Romania.

Synopsis

Situated at the epicenter of a drought-prone region, the village of Aagaswadi lies in the Satara district on the Mann Hills of Maharashtra, in western India. Frequent droughts and the lack of livelihood opportunities have compelled the young residents of the village to leave for the cities, abandoning their homes and leaving behind the elderly and children. With neither adequate drinking water nor irrigation facilities, Bhimrao decides to dig a well to sustain both himself and his farm out of desperation¹¹.

3.6.1 Exploring the documentary's Themes and Meanings

This documentary brings to the fore the manifold connected themes that reveal how the lives of Aagaswadi's residents have been deeply affected by environmental crisis and climate vulnerability. As farming becomes harder and opportunities disappear, the villagers' confrontation with ecological pressures has caused massive drought-induced migration with many young people fleeing to urban areas. This rural exodus immediately takes on a deeper meaning when one thinks that migration often introduces physical separation which inevitably modifies daily interactions; when individuals are forced to relocate, the intricate web of family bonds undergoes profound transformations, often resulting in possible role shifts to adapt to a new situation (Lifestyle Sustainability Directory, n.d., 2025, online).

¹¹ This brief synopsis was readapted from <https://alteff2020.eventive.org/>.

However, despite all these major obstacles, Bhimrao's decision to dig a well in response to despair symbolizes his sense of belonging as opposed to detachment. It is a clear gesture of hope and determination which reflects the importance of human agency, showing how his resilience turns out to be crucial in the struggle to protect his homeland.

Overall, the film brings out the significance of niche film festivals as beacons of diversity and cultural richness. Such festivals mainly cater to specific themes, providing a platform for filmmakers to amplify underrepresented voices as counter-narratives for marginalized groups or social issues that might otherwise remain unheard in the mainstream film industry. Therefore, tales about indigenous minorities, gender identity or restricted subcultures are of paramount importance to encourage understanding across diverse communities and promote inclusion in cinema (Dubai International Cine Carnival, 2024, online). Indeed, as an essential component of cinephilia, film festivals openly prioritize alternative cinema over commercially driven mainstream exhibitions (Uzzo, 2024), with special attention to accessibility.

More importantly, *Holbole* takes full advantage of his function as a vehicle going beyond simple storytelling to disseminate contents that must not remain unnoticed not to mention his ontological narration which encapsulates participatory stories bridging the personal and public domains to gain full perception of his role in the world (Somers & Gibson, 1994; Baker, 2006).



Figure 10: A frame from *VITS*.

Apart from language-related learning goals, the film was considered suitable for the task under this conceptual framework, aiming to raise awareness of marginalized content while emphasizing the importance of making it accessible to BVIPs through AD.

Despite its brief duration, the learners were asked to describe a short clip showing the opening scenes in which the village is introduced to the audience.

3.6.2 Text composition of *Aagaswadi – Village in the Sky*

The functional characteristics (Table 7) of the AVST refer to the excerpt presented to learners.

AVST: <i>Aagaswadi – Village in the Sky</i>		
Level	Parameter	Analysis
Extratextual	Intended text functions	Informative (primary) / Expressive (secondary), as realised through the visual introduction of the film
	Audience	General audience; in this study, learners producing AD for didactic purposes
	Medium	Documentary film excerpt (opening sequence) rendered through AD (written form)
	Addressees (sender/receiver)	Sender: filmmaker + AD describer; Receiver: documentary audience (intended), with learners acting as scriptwriters
	Motive	To introduce the environmental and social context of the village; didactically, to foster precise and context-aware description
Intratextual – Macrolevel	Genre	Documentary (observational / ethnographic)
	Text composition	Opening sequence structured as a gradual visual exposition of the setting, with no immediate narrative development
	Subject matter	Remote rural environment; initial glimpse of daily life and landscape
	Protagonists	The village and its inhabitants, presented collectively
Intratextual – Microlevel	Language function	Primarily referential: meaning conveyed through visual presentation of place and activity
	Language dimension	Observational, with no mediation
	Text focus	Spatial orientation and environmental context
	Specific visual elements	Mountainous terrain, village layout, natural surroundings, early signs of daily activity
	Filmic cues	Wide shots, slow camera movement, long takes, natural light
	Affective implications	Calm, remoteness, contemplation; gradual immersion rather than immediate engagement

Table 7: Functional characteristics of *VITS* opening scene (author's elaboration).

The primary function of the opening sequence is informative as it introduces the geographical and environmental context of the village, with particular emphasis on the effects of drought. At the same time, a subdued expressive dimension emerges through the visual composition of the landscape: the village is gradually exposed as a single entity, where the viewer is oriented through its geographical markers. The dry terrain, sparse vegetation, and slow rhythm of the shots contribute to a quiet sense of hardship, where the inhabitants appear as hushed, semi-static figures with minimal individual agency and almost folded into the landscape. In particular, this brings into focus both their living

conditions and the environmental constraints, hinting at a deeper reality that may be read as a subtle counter-narrative to more dominant representations of rural life.

On a macro-level, the description must provide clear spatial and contextual information to understand the village and its living conditions. Arguably, the main challenge for learners is to resist the urge to over-explain without slipping into vague and generic descriptions. At the micro-level, the referential precision typical of non-fiction AD (Cámara & Espasa, 2011) must not be seen as a limit to creativity; instead, learners are given freedom to interpret the most salient visual elements – such as the landscape, the layout of the village, and the people’s actions – clearly and coherently, provided that these remain grounded in the visual scene.

Indeed, the scene’s observational tone allows for a degree of creative expansion, where lexical variety plays a central role in the choice of nouns and verbs for the landscape, the people, and even minimal action. In professional AD, balancing referential precision and creativity would pose a significant challenge when describers are asked to settle for objective formulations, avoiding excessive expressivity. For the purpose of the study, learners are encouraged to pivot from neutral phrasing towards imaginative solutions, including – but not limited to – light figurative language or even poetic approaches, that would depart from conventional AD norms on the one hand, but open up more expressive and evocative descriptions, on the other.

Last, but absolutely not least, this clip is the only one including text-on-screen in the form of opening credits, whose core narrative function is to establish the initial framework through which the audience will interpret what follows. This introduces an additional challenge, as such elements may be omitted despite their narrative relevance, when they must, in fact, be integrated in the description and rendered either synchronously or asynchronously (Matamala & Orero, 2015). Here, the absence of particular time constraints gives learners sufficient silent gaps to insert the credits synchronously, by adding a word or explanation before it (*ibidem*); in the case of *VITS*, it might be: “*On a black screen, white text reads*”. The following table outlines the AD elements of the documentary:

Analytical dimension	Descriptive focus	Associated challenges
Narrative progression	Gradual exposition of the village and environmental context	Risk of over- or under-description
Pacing	Slow, extended shots, minimal action	Same as above
Focalisation	Subdued mood, landscape, village	Organizing information clearly; risk of vague descriptions

Analytical dimension	Descriptive focus	Associated challenges
Referential precision vs. creativity	Accurate description of visual elements combined with lexical variation	Combine information with expressivity
Text-on-screen	Integration of written information (credits)	Risk of omission
Lexical and stylistic realisation	Use of varied vocabulary and evocative language	Overly neutral renderings

Table 8: Descriptive dimensions and associated challenges in the opening sequences of *VITS*.

The AD elements will be discussed in Chapter 5, as in the two other clips.

3.7 Learners: the participants in the didactic experiment

Following the analysis of the three genres, we now turn our attention to the learners participating in the activity. The DAD task involved two groups of undergraduate learners in their third and fourth year of upper secondary school (aged 16-17), coming from two distinct institutions with different curricula (Table 9).

Participants	School	Duration	Attendance
Group A	IIS “Francesco Crispi” (Ribera - Agrigento)	2 days	Day 1: 21 learners Day 2: 9 learners
Group B	Liceo Linguistico Statale “Ninni Cassarà” (Palermo)	2 days	Day 1: 20 learners Day 2: 20 learners

Table 9: Participation details by school, including duration of the activity and daily attendance.

As shown above, group A initially included 21 learners in their third and fourth year of *Liceo Delle Scienze Umane* on the first day, with attendance dropping to only 9 of them on the second day due to other curricular activities. Conversely, group B consisted of 20 learners attending the third year of *Liceo Linguistico*¹² on both days. Despite coming from two different educational tracks, all the participants reportedly had a B2 level of English

¹² *Liceo delle Scienze Umane* and *Liceo Linguistico* are two types of Italian upper secondary school in the field of Humanities, but while the former focuses on human sciences such as psychology, pedagogy and anthropology, among the others, the latter is more oriented towards modern foreign languages.

with basic knowledge of general AVT practices but zero experience in them. The following figure refers to their understanding of AD, based on a quick introductory survey:

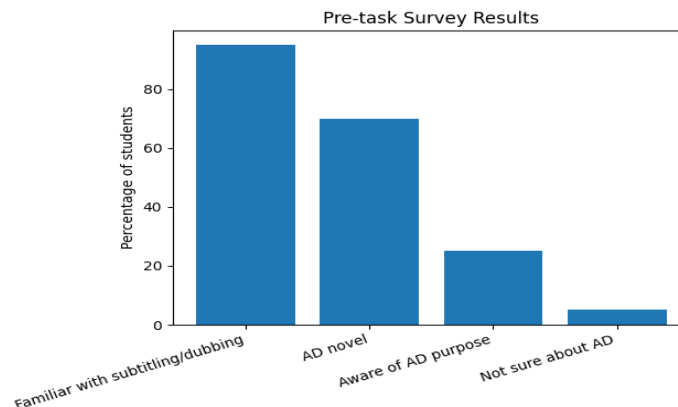


Figure 11: Percentage of the participants' knowledge of AVT and AD.

The survey revealed that while most of the participants (95%) were already familiar with subtitling and dubbing, AD remained relatively novel to them (70%), meaning that although they could easily recognize the main characteristics of the former two, their knowledge of the latter was limited, with only 30% having any prior awareness of its purpose or applications.

3.8 CADAD as a type of Didactic Audio Description

The task was designed to apply DAD as an innovative resource with significant potential for enhancing L2 proficiency, including cultural awareness, critical and creative-thinking skills. Following the emergence of creative practices within DAT in FLE contexts, creative-thinking skills were positioned as a key element for the task, based on the idea that they can be definitely taught (Talaván & Lertola, 2025). Such skills – also known as “little-c creativity”¹³ – involve everyday activities that can be cultivated through education if teachers help learners by encouraging exploration, generation and reflection on ideas (Kaufman & Glăveanu, 2021; Kaufman & Beghetto, 2009). Based on this perspective, the activities were organized around a combination of the models and lesson plans proposed

¹³ “Little-c” creativity is the opposite term of what Kaufman and Glăveanu call “Big-C” creativity, which refers to technology, literary or artistic works that require a high level of field-related knowledge (Kaufman and Glăveanu, 2021).

by Talaván *et al.* (2024) for an active use of creative DAT in FLE with the exception of Didactic Free Commentary (DFC):

- **Intersemiotic DAD**
- **Intersemiotic DAD plus intralingual L2>L2 (in the presence of dialogues)**
- **Creative DAD**

Intersemiotic DAD refers to the process of writing descriptions when no spoken dialogue is available; learners “are invited to [...] describe place and time, physical attributes of people (such as age, outfit, facial expressions body language, etc.), as well as their emotional state and actions” (*ibidem*). Conversely, Intersemiotic DAD plus Intralingual L2>L2 refers to the same activity in the presence of spoken interactions in L2, meaning that they will have to produce scripts in that language. Finally, Creative DAD requires learners to create a completely new script to elicit a particular effect (e.g., humor) or even produce their own videos (Navarrete, 2020).

For the purpose of the study, a hybrid approach to **Intersemiotic Creative DAD plus Intralingual** was deemed the most appropriate, resulting in a blended pattern that combines elements from each model, in which Creative DAD is largely defined by lexical choices. This approach inspired the formulation of the term **Creative Didactic Audio Description** (henceforth **CADAD**) which retains both intersemiotic and intralingual modes and was used with dialogues (*ES*) and without dialogues (*BE; VITS*), as shown in the following figure:

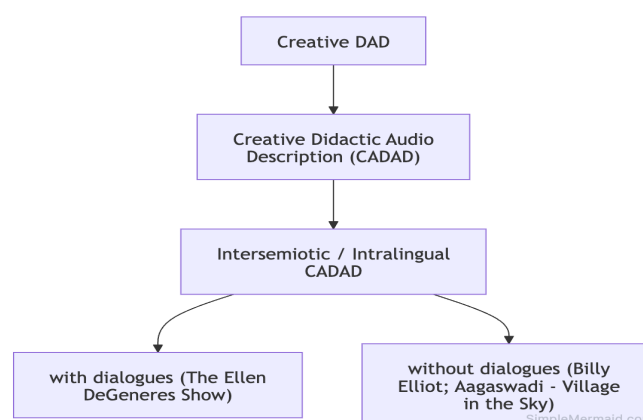


Figure 12: My elaboration of the use of CADAD in the task.

Following this model, the CADAD lesson plans were adapted to the three genres and organized as follows:

CADAD with no dialogues

CEFR level B2-C1

Video fragment: *BE*, acceptance letter scene;

Didactic AD: Creative intersemiotic (images > L2)

AD software: *Frazier - VideoToVoice*

Didactic functions: Improving vocabulary and writing production skills.

Aims: To practice creative rich language (adjectives related to emotions and feelings, figurative language); to experiment with interpretive/subjective AD styles; to detect cultural elements.

Structure:

Pre-task. “Blind” listen: Play video without showing images; learners guess what is happening on screen. Learners then watch the video twice and discuss whether their guess was right.

Script-writing They describe by following the given instructions. Learners can work either individually or in small groups

Post-task Prediction check for creative thinking: they predict the actual outcome of the letter. They watch the ending of the scene and check whether their guesses were right. Writing exercise with open questions. Correct.

CADAD with dialogues

CEFR level B2-C1

Video fragment: *ES*, Halloween costumes segment

Didactic AD: Creative intersemiotic (images > L2) plus intralingual (L2>L2)

AD software: *Frazier - VideoToVoice*

Didactic functions: Improving vocabulary and writing production skills.

Aims: To practice creative rich language (expressing action, physical appearance); to experiment with interpretive/subjective AD styles; to detect cultural elements.

Structure:

Pre-task “Blind” listen: Play video without showing images; learners guess what is happening on screen. Learners then watch the video twice and discuss whether their guess was right.

Script-writing They describe by following the given instructions. Learners can work either individually or in small groups.

Post-task: Do the listening exercise “A Rollercoaster of emotions” with Ellen’s monologue. Ask short comprehension questions. Show them the dialogue and discuss any cultural items.

They do the questionnaire. Correct.

CADAD with no dialogues

CEFR level B2-C1

Video fragment: *VITS*, opening scene

Didactic AD: Creative intersemiotic (images > L2)

AD software: *Frazier - VideoToVoice*

Didactic functions: Improving vocabulary and writing production skills.

Aims: To improve writing production skills. To practice creative rich language autonomously (action, adjectives, *etc.*). to detect any cultural elements.

Structure:

Pre-task - They have already done the “blind listen” during the IP.

Script-writing They describe by following the given instructions. learners can work either individually or in small groups.

Post-task Brainstorming. Correct.

Extra Activity: Roleplay activity with the learners acting as audio describers for a clip of their choice (*BE, ES, VITS*).

(My adaptation of the lessons plans proposed by Talaván *et al.* (2024))

During the CADAD activities, which will be further discussed in the next section, writing production in L2 is also expected to have positive effects on learners’ idiomaticity, due to the substantial reduction of any possible involvement of L1 (Pensabene, 2023:59,

cf. Bausells-Espín, 2022), although the process of mental translation (be it intralingual or intersemiotic) is very likely to be influenced by their native language at this level.

Talaván *et al.* (2024) also provide language instructors with a couple of useful extra prompts, such as the addition of original scripts or the inclusion of a list of intercultural elements as scaffolding options. Regarding these suggestions, the integration of original scripts was deliberately omitted to reduce potential lexical interference; conversely, CSRs were discussed in class and the learners were accordingly encouraged to detect any relevant example that they could incorporate in their scripts.

Instead of producing a completely different script, the learners were expected to express their own creative intervention through the use of subjective and interpretive solutions resulting from both linguistic and stylistic strategies. The former mainly relied on evaluative lexical carriers such as adjectives, verb choice, intensity markers and lexical selection for CSRs, if we think that the representation of culture-bound elements is not simply a matter of semantic transfer, but rather occurs at a lexical level in the first place and is therefore encouraged by lexical accuracy (Rizzo & Spinzi, 2023:424). The latter, instead, is closely related to the emergence of creative/alternative practices in MA promoting subjectivity as a strategy to reconsider the conventional assumption that objective descriptions should be the norm. This means that, although the participants were informed about the conventional rule of factual AD as an accessibility mode, a subjective style was also encouraged and, to prevent AD from resembling a science experiment (Romero Fresco & Dangerfield, 2022), they were provided with a short list of figurative expressions for the *BE* script and were also given the freedom to use evaluative words or to vary sentence length and rhythm as carriers of subjectivity.

To some extent, lexical richness alone is arguably the main catalyst for subjectivity throughout the descriptive process, where the use of qualifying adjectives and adverbs of manner largely depends on the author's perspective (Rizzo & Spinzi, 2023), especially when describers look into the characters' psyche and the cultural specifics of the original product (*ibidem*), as in *BE*, *ES* and *VITS*.

Now, one could argue that pushing first-time describers towards alternative uses with subjective stances might generate confusion about the true purpose of AD, making it less suitable for a blind audience and more applicable to a mere stylistic exercise. This is true to a certain extent, and elaborating on this idea, we may refer to the AENOR (*Asociación Española de Normalización y Certificación*)'s requirements for quality in AD,

where objectivity is strongly encouraged “to protect persons with visual impairment from excessive interpretations, paternalistic descriptions or exposure to the audio-describers’ personal opinions” (cf. Bardini, 2022:105). Instead, a neutral approach accordingly relates to AD that uses language describing visually observable content “where the emphasis is on what’s manifest [...] and whose primary aim is to inform or describe” (Kleege, 2016, cf. Soler Gallego & Luque Colmenero, 2023).

On the other hand, if we look at the core aims of AD rather than its educational functions, this choice aligns closely with scholarly discussions advocating a more interpretive and artistic approach to AD that revolves around expressive lexical items for improving narrative engagement and immersion (Holsanova, 2016; Marra, 2023), or subjective language for handling implicit elements, *i.e.*, characters’ mental states (Geerinck & Vercauteren, 2020), to cite a few.

Last but not least, it is important to point out that this study was originally structured to include intentional methodological deviations, making CADAD a genuinely creative exercise. The main implication behind this is that the script-writing activity fully draws on an accessibility modality and becomes an L2 writing enhancement exercise which deliberately challenges conventional practices. This choice chimes with the functionalist framework, where the pedagogical *Skopos* of AD – the development of the learner's linguistic and expressive repertoire – redefines the efficacy of the intersemiotic transfer. By prioritizing communicative impact over rigid constraints, the task allows for a flexible application of (audiovisual) translation norms, thus ensuring that AD serves primarily as a vehicle for linguistic growth and creative exploration.

3.9 CADAD Task: Design and Implementation

In general, research into DAT is either part of longitudinal or wider-scale projects facilitating in-depth investigation and methodological replicability. Conversely, the purpose of this workshop was to raise awareness on AD as a complementary FLE tool, running over two separate seven-hour sessions for group A and B, with all the activities condensed to span both days in a fully immersive experience.

As previously discussed, the participants were asked to write three CADAD scripts in English for three short excerpts from *BE*, *ES* and *VITS* with the help of a thesaurus if

needed. The following table (10) gives a general overview of the course evolution over the two-day sessions:

	Day 1		Day 2	
Group A and B	<i>BE</i>	<i>ES</i>	<i>VITS</i>	
	Pre-task Phase (PP1)	Pre-task Phase (PP2)	Pre-task Phase (PP3)	
Introductory Phase (IP)	Script-writing Phase (SWP1)	Script-writing Phase (SWP2)	Script-writing Phase (SWP3)	
	Post-task Phase (PTP1)	Post-task Phase (PTP2)	Post-task Phase (PTP3)	

Table 10: Structure of the CADAD workshop.

As illustrated above, the activity was divided into four main phases: 1) Introductory Phase (IP); 2) Pre-task Phase (PP1, PP2, PP3); 3) Script-writing Phase (SWP1, SWP2, SWP3), and 4) Post-task Phase (PTP1, PTP2, PTP3), but while the IP only took place at the beginning of the course, the other three were adapted to and repeated for each CADAD assignment.

3.9.1 Introductory Phase

The IP began with an oral pre-assessment knowledge check about their familiarity with AVT and AD (see Section 3.7, Fig. 11) after which the learners watched short sample videos showing interlingual, intralingual subtitling and AD. Of the 30% familiar with AD, the majority thought of it as description of movement, while only one respondent associated it with transcription. At this point, the focus quickly shifted towards a more detailed explanation of AD and CADAD purposes for accessibility and FLE and, once they had a better grasp of the topic, they were asked how they believed descriptions should be structured. Interestingly, the overwhelming majority felt that AD should be clear, basic and brisk-paced.

After finally acknowledging its strategic importance for MA and inclusion and the idea that the People with Sight Loss (PSL) should experience and enjoy a visual product as

closely as possible to a sighted audience (Perego & Taylor, 2022), they listened to the excerpt from *VITS* without access to visual content, relying only on the auditory channel. This immediately shed light on their initial bias for basic descriptions while also helping them to build up their mental representation of the setting.

Interestingly, while over half of the learners imagined the sound of either a plane's turbine or a car engine, only a minority of them visualized forest and savannah wildlife, pristine or prairie landscapes which were actually not far off from the original scene. This moment was crucial in cementing the idea that describers should take an audience-centered perspective, creating intelligible and stimulating content through complex vocabulary to convey an engaging message to the listener (Taylor, 2022), regardless of its readability and listenability. Building on this perspective, the learners were introduced to their CADAD task before receiving the short list of targeted guidelines reported below:



Audio
Description

Tips for your AD task

- Describe the visual elements that are essential for understanding the message of the video. Imagine you are describing it to someone who cannot see it. Avoid mentioning details that are already evident from the audio.
- Describe who is on screen, where the scene takes place, how people look and dress, their expressions and movements, and any sounds that are not clearly identifiable.
- Use **accurate**, **vivid** and appropriate language and avoid the term “We See”.
- Write description in present tense, active voice, and third-person narrative style.

- Synchrony: “The AD script must appear synchronized with the corresponding images (not appearing before, not staying longer) and never overlap the dialogues, if any” (Talaván *et al.*, 2024).

Source: *makethingsaccessible.com*

Creative variations as language build-up activities:

Feel free to replace subjective, neutral descriptions with more creative solutions as long as vivid, accurate language is retained. In order to do so, focus on the use of varied adjectives and specific verbs to describe characters, settings and action to avoid generalizing. The use of figurative language with metaphors, personification and similes is also encouraged.

This will make descriptions more imaginative and effective.

Examples of figurative language:

- *The table’s legs are **like** fluted classical columns* (Spinzi, 2019);
- *That movie was a **rollercoaster ride** of emotions;*
- *The carer was **as gentle as a dove**;*
- *A **wave** of terror washed over him;*
- *I have been **a busy bee** today.*

Examples of vivid vocabulary for creative language

Ex. Kate is angry

Kate clenches her fists and stomps out of the room

Verbs

Suggested synonyms with different shades of meanings:

Walk

stride, stroll, hike, stomp, wander, amble, meander, march, strut;

Get

earn, win, grab, clasp, fetch, capture, secure;

Make

craft, assemble, prepare, fashion, create, manufacture;

Look

gaze, peer, inspect, glance, stare, eye, scrutinize, watch, observe;

Describing people: adjectives connected to feelings and emotions

Positive meanings: elated, ecstatic, contented, jubilant, euphoric, relieved, serene, thrilled, stunned, buoyant, joyful;

Negative meanings: enraged, disgruntled, anguished, indignant, mournful, regretful, fearful, devastated, appalled, miserable, bewildered, despondent.

CSRs

Culture-specific items are concepts that are specific for a certain culture. These concepts can refer to [...] flora, fauna, food, clothes, housing, work, leisure, politics, law, and religion among others. (Persson, 2015). The describer must detect any relevant cultural reference and adapt it in the text.

AD script-writing tasks:**1) *Billy Elliot* (comedy-drama)**Pre-task:

Listen to the video without images and try to guess what is happening on screen.
After watching the clip twice discuss whether your impressions were right.

Script-writing:

Write the AD script using adjectives related to emotions and feelings.
Try to detect any cultural elements.

You may want to incorporate figurative language in your descriptions; feel free to use the following expressions in your description:

1. **Tied up in knots**
2. **On edge**
3. **Jitters**
4. **Fretting**
5. **Nervous wreck**
6. **On pins and needles**

Post-task:

Do the writing exercise and follow your teacher's instructions for further task-related activities.

2) *The Ellen DeGeneres Show* (talk-show)

Pre-task:

Listen to the video without images and try to guess what is happening on screen. After watching the clip twice discuss whether your impressions were right.

Script-writing:

Write the AD script while focusing on action (experiment with verbs).

Try to detect any cultural elements.

- The children's fancy dresses are two examples of US cultural items. If you don't know the meaning behind these costumes, you may do some research and decide how to introduce them to the audience in the AD script. Can you spot any other cultural items?
- Post-task:
- Do the listening and quiz and follow your teacher's instructions for further task-related activities.

1) *Aagaswadi – Village in the Sky* (documentary)

Pre-task:

Listen to the video without images and try to guess what is happening on screen.
After watching the clip twice discuss whether your impressions were right.

Script-writing:

This scene depicts a more static situation; write the AD and feel free to put emphasis on some features from the previous tasks (e.g., verbs, adjectives, figurative language, cultural elements).

Post-task:

Follow your teacher's instructions for any other task-related activities.

(My elaboration and adaptation)

The reference to CSRs in the instructions sparked the learners' curiosity in learning more, which started a general class discussion with a brief oral survey.

However, they did not begin their CADAD activities until they were provided with necessary training in the use of *Frazier® - VideoToVoice* (henceforth also *VtV*), the AD web-based software for creating their scripts, which was the final step of the first part.

3.9.2 *Billy Elliot*: PP1, SWP1, PTP1

By the time the course transitioned out of the IP and into the first CADAD task, the learners were prepared enough to take an active role in the script-writing process.

The *BE* three-step pattern began with a “blind listen” exercise based on an audio-only preview of the clip. Because all the learners were unfamiliar with the scene, both the lack of visuals and the absence of dialogue made comprehension challenging; however, the auditory cues were still sufficient for them to perceive the high tension of the moment. When asked what had struck them most, the learners reported feeling a gradual build-up of anxiety triggered by the emotional piano melody and later intensified by what they believed was Billy's shallow breathing.

When they finally watched the scene, the reactions were varied: while some felt their initial predictions were confirmed, others were surprised by the scene's visual intensity, calling it a "dense moment" and "a nerve-wrecking situation".

At this point, they were left free to start their descriptions on *VtV*, meaning that they were allowed to play the clip as many times as necessary to capture any relevant details on the platform. As first-time authors of audiovisual content, the participants also had the option to join forces in small groups and exchange opinions during the creative process, while their activity was being monitored from the main control station. This was one of the most revealing aspects of the study: keeping track of their workflow during all the SWPs offered a significant insight into their editing process, while also giving quick real-time suggestions via *VtV*. Even though the wording structure selected during their drafting phase sometimes appeared more appropriate than their final versions, immediate feedback was limited to any major grammatical mistakes, with the learners maintaining full agency on their lexical choices.

The CADAD clip stopped a few seconds before they could learn whether Billy had been accepted into the Royal Ballet School, so when all the descriptions had been completed and double-checked, they were genuinely curious to find out about the actual outcome of the letter, which was revealed, in fact, during PTP1. During this phase, once again, the learners' creative thinking was put to the test with a prediction check: they were given a two-item open questionnaire and asked to answer the first question focusing on their expectations. After completing their responses, they were ultimately invited to watch the continuation of the scene, in which Billy's father wrenches the door open with all the other family members clustering right behind him to hear Billy murmur in tears: "*I got in*". The scene ends with Jackie dashing towards Everington Miners Club and blurting out the exciting news to his co-workers.

At this point, they were asked to answer the second question prompt – as illustrated below – requesting them to provide an objective description of the event, comparing it with their initial predictions.

BE - Post-Task Questionnaire as per author's own elaboration

1) Your teacher is going to show you the second part of the acceptance letter scene. Before watching the clip, imagine what might happen next and write your predictions below as if you were actually watching the scene. Focus on the characters' actions and emotions.

2) After watching the second part of the scene, rewrite the scene objectively with factual sequences. Compare and discuss with the class.

Interestingly, the majority (around 85%) predicted that Billy would be accepted into the ballet school, though they were expecting different emotional reactions – ranging from being “*stunned*”, “*uncertain*”, or having “*mixed emotions*” – due to his family’s financial situation. Conversely, a limited group (around 15%) believed that Billy would be “*devastated*” following the rejection from the school. Curiously, the learners quickly picked up on the tension between Billy and Jackie, so it was small wonder that many of their predictions were predominantly focused on his father’s reactions. A distinct 35% expected Jackie to be “*running to the miner’s center*”, “*running quicker than light to tell relatives*”, or “*running straight to Mrs Wilkinson to share the great news*”.

Both groups were positively impressed by Jackie’s evolution from a cold, intimidating father to a proud man, as he was the one actually leading the celebration moment.

3.9.3 The Ellen DeGeneres Show: PP2, SWP2, PTP2

The second CADAD task with dialogues began with a general presentation of *ES* as a genre requiring a different approach. Notably, although it was never aired on Italian mainstream TV channels, the show had widespread popularity and resonated quite well with the learners. Before then, they had randomly watched some segments featuring celebrity interviews, but the Halloween episode was reportedly new to them.

PP2 was introduced with the same “blind listen” moment, during which the learners could hear Ellen’s humorous exchange with her young guests. Although they had a clear understanding of what was happening on screen, they had no visual access to the costumes, so, when asked to describe their mental representation, a significant majority (55%) imagined Clint Eastwood appearing in his Mexican-style poncho, while some of them candidly admitted not having heard of him until then. Interestingly, as there was no firsthand vision of the Chair costume, when Ellen said that the two costumes were to be seen in tandem, they believed the second girl impersonated a lady in a typical nineteenth-century

Western gown, a visual trope reminiscent of female characters appearing in *Spaghetti Western* films¹⁴. This visual gap was further complicated by the third costume, but while the laugh track suggested a humorous impact, the Magic Mike character only sounded vaguely familiar to them, with the majority struggling to recognize his distinctive traits. This is hardly surprising, if we think that: a) the learners were still too young when the film peaked in 2012, and b) it did not become a household name like other franchises mainly for its provocative nature. However, the “blind listen” exercise had piqued the participants' curiosity to such an extent that, when they finally watched the video, they were captivated by the actual costumes – a moment that shifted the focus to a more detailed discussion on the importance of culture-bound elements.

When they started SWP2, they were completely absorbed in the editing process. Admittedly, this session was comparatively more demanding: while *BE* primarily focused on the characters' psychological traits, the *ES* segment placed equal importance on action and physical appearance. This posed a considerable challenge for the learners as describers “on the go”, who had to decide whether to prioritize the children's movement or their costumes without watching the entire episode (Jankowska, 2021). At the end of SWP2 and contrary to expectations, both groups expressed general appreciation for the segment, indicating the fast-paced setting as the most challenging element to manage.

PTP2 temporarily moved away from the CADAD towards a more language focused task reviewing emotions through adjective phrases (*AdjP*). Here the learners were requested to answer a short multiple-choice questionnaire based on an excerpt of one of Ellen's interventions from a different episode, called “*A Rollercoaster of Emotions*”. In this specific segment, Ellen discusses a scientific study from Ohio State University about compound emotions expressed through a combination of feelings such as “happily surprised” or “fearfully disgusted”, which translate into a combination of adjective + adverb from a grammatical perspective. As shown below, the questionnaire mainly served as an extra vocabulary exercise asking the learners to provide synonyms for the various emotions included in, but not limited to, the segment.

¹⁴ *Spaghetti Western* refers to a broad subgenre of Western films that emerged in the mid-1960s, distinguished by their Italian production, direction, and filming locations (primarily the deserts of Almería, Spain, and studios in Rome). Brought to prominence by director Sergio Leone and composers like Ennio Morricone, the genre reconfigured the traditional, North American Western into a gritty, stylized landscape defined by anti-heroes, stark violence, operatic music, and distinct visual codes (e.g., Eastwood's popularized poncho, the *cigarillo* chewed or held stationary between the teeth, extreme close-ups, just to name a few.)

ES – A Rollercoaster of Emotions

After watching the clip, complete the following tasks.

1) Find an adjective that might replace the expression "happily surprised".

- A) thrilled
- B) shocked

2) Choose the adjective that might replace the expression "sadly angry".

- A) miserable
- B) sullen

3) Choose an adjective that might replace the expression "fearfully disgusted".

- A) horrified
- B) disturbed
- C) shattered

4) The following adjectives all have negative meanings except one. Choose the odd one out.

- A) appalled
- B) upset
- C) annoyed
- D) elated

<https://docs.google.com/forms/d/1DWyPfxlMCrnsOmYmtQ3Prcf3FEAzGZXNrl6RZwP0gal/edit> 1/225/05/24

The learners watched the video twice – with and without subtitles – and were also provided with the script of Ellen's introductory speech; however, they were invited to complete the *Google Docs*-generated questionnaire at the end of the sessions and submit their answers remotely to allow more time for the remaining sessions of the workshop.

3.9.4 Aagaswadi – *Village in the Sky*: PP3, SWP3, PTP3

For the final CADAD task, the learners were introduced to another silent clip. This redirected their focus toward a static scene, which facilitated decision-making on the one hand, while encouraging a more detailed narrative structure, on the other. It should be noted that the “blind listen” for the *VITS* clip was one of the activities of the IP, so PP3 was deliberately deleted. This significantly reduced the workload, allowing learners to focus exclusively on SWP3.

As already shown in Section 3.2, the last CADAD exercise was essentially a consolidation of the previous ones. This gave the learners the opportunity to prioritize any features of their choice, meaning that they could decide what should be emphasized without task-based instructions.

Similarly, PTP3 was omitted due to time constraints and accordingly replaced with a role play activity, during which some participants acted out as professional describers reading out their scripts. A small group volunteered to pair up and work on *BE*, taking turns as describers while their partner sat facing away from the screen to simulate the experience of a blind person. Needless to say, everyone already knew the scenes well, so this simulation mainly served as a final wrap-up for consolidating the purposes of AD within a more relaxed environment.

3.10 *Frazier*® - *VideoToVoice* as a Pedagogical tool for CADAD

The integration of AVT-oriented methodologies within language learning environments has been greatly accelerated by the emergence of digital technologies. Much like other DAT modes, DAD can be adapted to meet the demands of new media forms preserving multimodal approaches (Baranova & Duoblienè, 2019; Brescia-Zapata & McDonagh, 2024). This process falls under what is known as Digital Multimodal Composing (DMC), which not only enhances linguistic competences, but also nurtures creativity and critical thinking, as digital technologies continue to reshape educational practices (Li, 2021; O'Halloran *et al.*, 2017; Talaván *et al.*, 2024).

Notably, such innovative approaches prompt L2 learners to explore new identities within DMC (Li, 2021, cf. Zhang & Yu, 2024) which, in the case of CADAD, sees learners

stepping into a technology-driven authorial identity. Building on this perspective, their experience as describers can be significantly optimized by AD-specific tools that can potentially develop learning through acquisition and collaboration (Mazur & Chmiel, 2021).

Indeed, the advantages of using such tools are twofold: pedagogically, they grant learners full agency as creative authors and editors over the multimodal process while practicing their L2 within a simulated professional scenario; practically, such tools handle the time-coding of AD units automatically (Minutella, 2022), making the AD process faster and smoother. Alternatively, some interesting suggestions also come from interactive storytelling platforms used for creative immersive approaches with accessibility purposes (Brescia-Zapata & McDonagh, 2024).

The CADAD scripts were produced using *VtV*, a German web-based platform that was singled out for its intuitive interface among a series of AD software, such as *LiveDescribe*, *Anglatècnic*, or *AD Author*, to name a few. First launched in 2017, this cloud-based AD tool uses Text-to-Speech (TTS) synthesis, generating spoken versions of AD units known as TTS AD (Schaeffer-Lacroix & Schulz, 2021; Szarkowska, 2011, cf. Kurch, 2024:422). As shown in figure 13, the main interface immediately comes across as simple and intuitive for the editing process, with a timeline strip at the bottom and the orange marker showing the current position in the video.

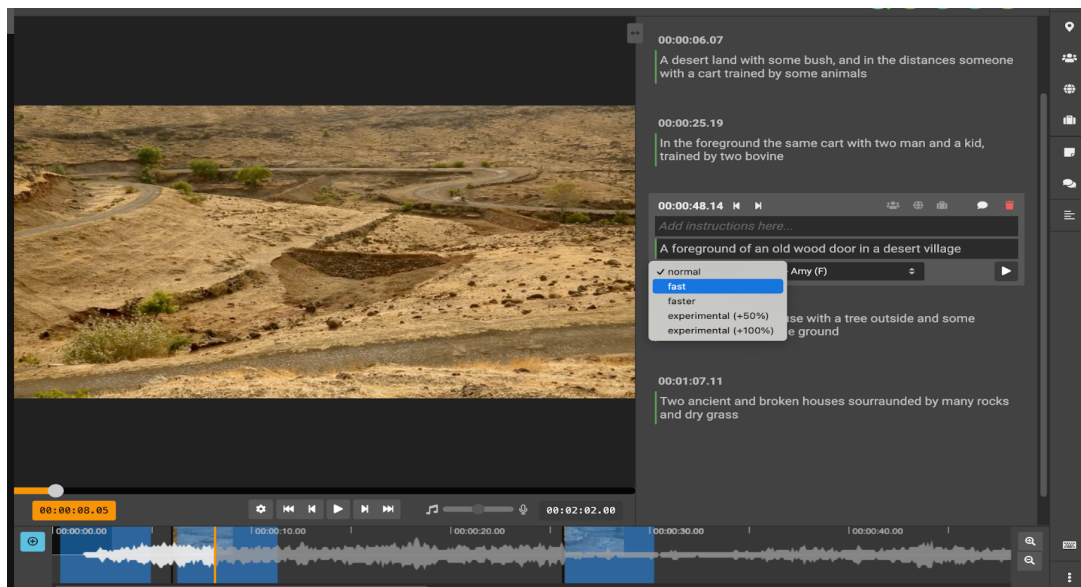
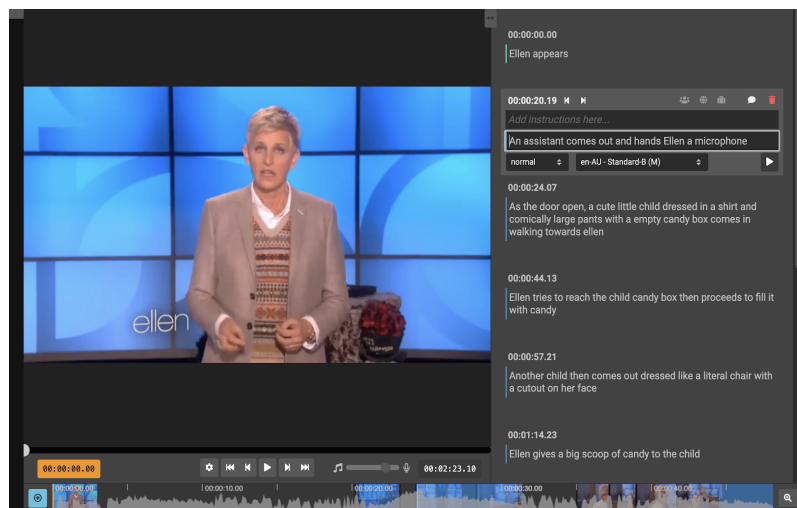
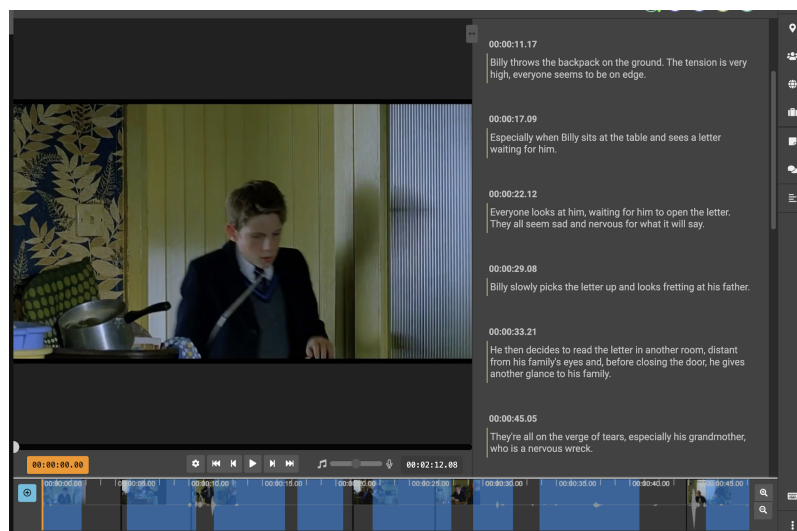


Figure 13: *VtV*'s main interface showing described segments and the editable text box. (source:videotovoice.com)

More importantly, using *VtV* in class allowed the learners to benefit from an all-in-one “Production Suite” in the browser, often called *The Frazier Loop*. This refers to a concrete series of keyboard shortcuts to create a description (videotovoice.com, online) – e.g., Ent + writing AD + Esc – that automatically captures the timecode while opening a text box and using TTS to read the description back to the user instantly. Simultaneously, a blue bar with the described portions appears on the timeline below the video signaling the text duration. During the CADAD tasks, the editing process was also supported by the “ducking function”, which automatically lowered the volume of the main soundtrack whenever AD was heard (Schaeffer-Lacroix & Schulz, 2021).



Figures 14 and 15: Examples of descriptions for *Billy Elliot* and *The Ellen DeGeneres Show*.

However, the learners particularly appreciated the TTS function and, even though the length of the passages was calculated automatically, “they could slightly adjust it by selecting the individual speed of the passage and even the voice reading it” (*ibidem*).

Among the suggested options, they could also include brief comments, create a terminology database with recurring words or names, called “snippets”, as well as insert “markers”. The latter serve as temporal cues indicating the precise moments when each segment of the description should begin and end, which helps maintain a natural narrative flow between dialogues. The learners could decide to add them as small bookmarks on the timeline marking a position in the video, tag the first appearance of a person or label a scene to return to for later analysis (videotovoice.com, online). For this reason, they were encouraged to use markers to locate and adjust the AD segments, knowing that this would significantly streamline the editing work.

There were initial concerns about the use of the platform, but doubts were soon mitigated by its intuitive design and comprehensive functions. Taken together, these functions turned out to be a major advantage for the entire CADAD process, as they significantly reduced the technical issues that usually slow learners down, making it easier for them to stay focused both on the writing and editing process.

Ultimately, we may say that working on VtV was about more than just using software; it was a way to bring the “digital innovations” into the classroom. When we move learners into a professional, cloud-based space, we are showing them the modern technological workflows they will come upon in the industry. Apart from the fact that the platform, for its part, can really slim down the technical synchronization of the AVST and AVTT, the use of a dedicated AD interface pushes the whole process beyond a simple script-writing exercise and gives learners a taste of the real deal, that is the true impact of technology on AVT.

Chapter 4 – A Corpus-based Analysis: Assessing CADAD through *AntConc* and ADLAB Guidelines

This chapter focuses on a quantitative analysis to evaluate the emerging lexical priorities from the CADAD scripts according to two paradigms: 1) a frequency-based analysis with a software-aided examination of empirical data; 2) a more detailed observation of the learners' lexical choices as per ADLAB guidelines.

The final section delineates the assessment framework, adapting CEFR/CV (2020) B2-C1 descriptors to intralingual AD, focusing exclusively on the script-writing task, where AD is both a linguistic mediation strategy and a creative production method.

4.1 Corpus Analysis: *AntConc*

The first step towards a quantitative approach involved a frequency-based observation to classify measurable lexical dispersion across the texts, which was conducted using Laurence Anthony's *AntConc* software (4.3.1 *MacOS 10/11 – Silicon* 2024 version). Laurence himself identifies *AntConc* as “a freeware, multi-platform, multi-purpose corpus analysis toolkit [...] featuring a comprehensive set of tools including a powerful concordancer, word and keyword frequency generators, tools for cluster and lexical bundle analysis, and a word distribution plot” (Anthony, 2004:7). The main reason for this choice is that *AntConc* was implemented for specific use in the classroom following the emergence of a “corpus revolution” phenomenon which seems to have been urging learners to explore corpora to investigate for themselves the use of language in target contexts (*ibidem*).

Another reason is that *AntConc* acts directly on the raw text, storing results in its active memory, which makes it appropriate for processing modest-sized corpora (Anthony, 2006:284). Nonetheless, corpus linguistics generally acknowledges that increasing the size of a corpus significantly reduces the likelihood that rare linguistic features will reflect a near-zero frequency, as “the corpus serves not to *test* a linguistic model but to *create* a linguistic model [...]” (Anthony, 2013:146 a critical look, etc.); therefore, “if the corpus is small it can only provide a small window on the language phenomenon under investigation” (*ibidem*).

Conversely, this study excluded the learners from the analytical process and was intentionally conducted on a small dataset to test the use of creative language in response to task prompts, seeing that the value of a corpus depends primarily on the type of

information that can be drawn from it, rather than its size (*ibidem*). Additionally, the corpus was used for a frequency-based examination, since extending the analysis to a semantic deconstruction may have introduced significant technical limitations, not to mention that *AntConc* primarily focuses on lexical patterns and recurring structures, with no integration whatsoever of ontology-based methods or conceptual categorization. Here, the combination of *AntConc*'s intended purpose and its intuitive interface made it a legitimate instrument for corpus segmentation, and, more importantly, while a detailed description is beyond the scope of this study, a brief outline of its key features will be necessary to see how the lexical data were collected, processed, and analyzed (Fig. 16).

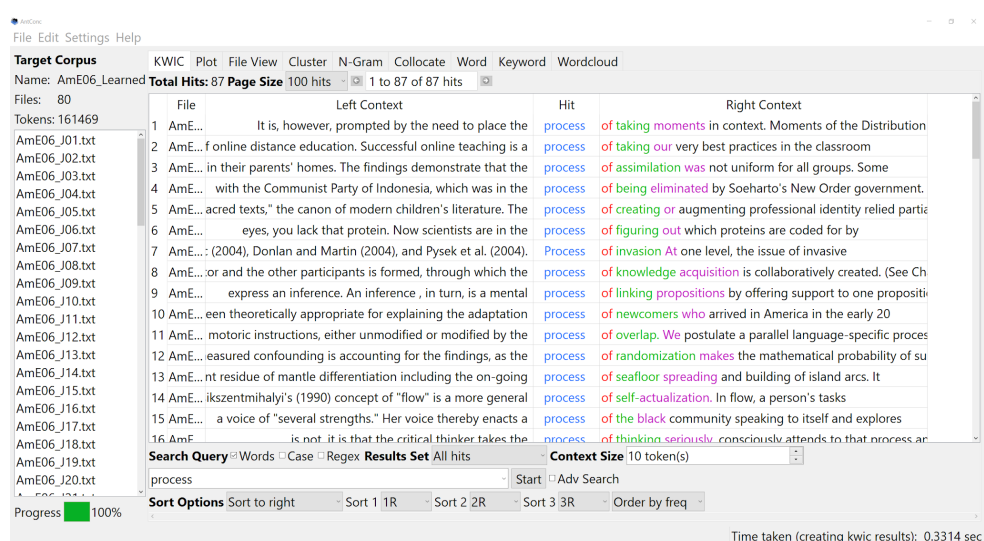


Figure 16: *AntConc* 4.3.1 main interface (laurenceanthony.net).

The image above shows its main interface with a sample text for a clearer understanding of its layout and practical use in corpus-based studies. Although there may be some slight differences in the various updated versions that have been released over the years, the functions of its operational tools have remained substantially unaltered. The left-hand column displays the number of uploaded files and the token count, whereas the top section shows the tool windows – with concordance results for the word “*process*”, also referred to as **hit** – which can be summarized as follows:

- **Concordance Tool**

It shows output results in a “KWIC” (KeyWord In Context) format to see how words and phrases are used in a corpus.

- **Concordance Plot Tool**

Here, outputs are displayed as a “barcode” format to detect the exact position of search results in target texts.

- **File View Tool**

It is used to see individual files, offering a detailed overview of the results generated in other tools.

- **Cluster/N-Grams**

The former detects clusters based on the search condition, summarizing the results generated in the **KWIC Tool** and the **Concordance Plot Tool**, while the latter scans the corpus to identify clusters of a specified “N” length (e.g., one-word two-word clusters, *etc.*).

- **Collocates**

It helps to display the collocates of a term to facilitate the analysis of non-sequential patterns.

- **Word List**

This tool counts all the words to find the most frequent single units in a corpus.

- **Keyword List** This tool finds the unusually frequent (or infrequent words) in the corpus, in comparison with a reference corpus¹⁵.

This brief overview provides a clear outline of the basic functions that are generally useful for breaking down a small-scale corpus, even though certain features – such as the **Concordance Plot Tool** and the **Keyword List** – were not applied in this study, whereas the use of the **Collocates** input was minimal. More specifically, the former were not relevant for this analysis, meaning that the exact position of the lexical units within a text did not appear to affect the research results, nor was a comparison with another reference corpus required. The latter, instead, was suitable for validating the concordances of the most frequently occurring single-word units when needed.

Having illustrated the primary features and functionalities of *AntConc*, the following section analyses the lexical units associated with each thematic CADAD subset in relation to the targeted results.

¹⁵source: *AntConc* Help File Version (Windows, Macintosh OS X, and Linux); Build 3.4.3.

4.2 Quantitative Analysis: measuring lexical frequency in CADAD Scripts

As explained in chapter 3, after completing their intensive two-day workshops, group A and group B had produced six sets of CADAD in total (3 genre-specific sets per group), with each group working on the same audiovisual content.

To proceed with the analysis of high-frequency lexical occurrences, table 11 shows how the six sets of scripts (henceforth also referred to as “texts”) were organized into three larger corpora; each corpus includes two clusters of scripts – one from Group A and one from Group B – corresponding to the three different clips.

Participants	School	Duration	Attendance	Audiovisual Content	AD Script Sets
Group A	IIS “ <i>Francesco Crispi</i> ”	2 days	Day 1: 21 learners Day 2: 9 learners	1. <i>Billy Elliot</i> (02:12 - drama) 2. <i>The Ellen DeGeneres Show</i> (02:23 - talk show) 3. <i>Aagaswadi - Village in the Sky</i> (02:02 - short film)	3 sets of CADAD (1 per genre)
Group B	Liceo Linguistico Statale “ <i>Ninni Cassarà</i> ”	2 days	Day 1 20 learners Day 2 20 learners	1. <i>Billy Elliot</i> (02:12 - drama) 2. <i>The Ellen DeGeneres Show</i> (02:23 - talk show) 3. <i>Aagaswadi - Village in the Sky</i> (02:02 - short film)	3 sets of CADAD (1 per genre)

Table 11: The CADAD scripts that make up the corpus.

It is worth noting that, although the majority of participants made effective use of the given time to complete the tasks, the inspection of individual texts revealed some instances of incomplete or entirely missing descriptions, as exemplified in the table below:

Participants		Corpus Type	
Billy Elliot		The Ellen DeGeneres Show	Aagaswadi (Village in the Sky)
Group A	Incomplete: 3	Missing: -	Incomplete: 5

	Missing: 3	Missing: -	Missing: -
Group B	Incomplete: 3	Incomplete: 12	Incomplete: 10
	Missing: -	Missing: -	

Table 12: Incomplete and missing descriptions across the three clips.

The table shows a considerable number of incomplete, fragmented descriptions (35 in total), due to the fact that putting instructions into practice was reportedly perceived as very demanding by some participants, even though the textual information provided was sufficient for the purpose of the analysis. Conversely, the limited number of missing texts can be explained by the learners' collaborative approach, as some of them decided to join forces and pair(ed) up to produce a single text.

It must be remembered that, according to the first paradigm, the analysis of frequency rates did not aim to assess the descriptions in terms of creativity, or lexical accuracy, which will be discussed further on. Conversely, focus was on lexical bundles with the highest aggregate frequency, broadly classified in line with the main foundational principles of AD scriptwriting (WHAT, HOW, HOW MUCH, WHEN), which were modified as follows for the sake of the analysis:

- a) The Spatio-Temporal Unity: I considered the spatial and temporal dimensions as a single, integrated unit;
- b) Qualitative Priority: Although quantitative metrics give a surface-level view of descriptive density, quantity (how much) was deliberately omitted from the data analysis to focus on functional saliency. In a pedagogical framework, the quality of a script depends less on its lexical volume and more on the student's ability to select high-value descriptors.
- c) The "WHO" Pillar: I added WHO as an extra foundational pillar alongside the traditional parameters, because character identification is the starting point for the description. In this way, the resulting "who" does "what" and "how" triad serves both as a prerequisite for intermodal coherence and a catalyst for creative AD.

That said, the first part of the analysis was based on WHO, WHAT, WHERE, WHEN to maintain a clear and logical structure. Following these methodological clarifications, a

close inspection of the two largest thematic datasets, namely *BE* and *VITS*, revealed different outcomes in terms of highly recurring lexical units: the former was characterized by a clumpy distribution of words, whereas the latter exhibited a more homogeneous dispersion. Because of this, a minimum threshold of 5 occurrences was established to filter the data, below which items were to be excluded from the analysis. Later, it was decided that all occurrences would be included, regardless of their frequency rate. However, the same search criteria for retrieving raw counts and relative frequencies were applied in all cases to assess how common or uncommon a word might be (Carradini & Swarts, 2023:88).

The frequency-sorted lists below of both single units and lexical bundles were extracted from the total number of tokens – with the exception of function words – and may contain grammar or spelling mistakes, as they appeared in the learners’ original submissions. All the lexical occurrences were ranked by: a) overall frequency of single units; b) frequency of units and patterns associated with task-specific vocabulary to support the creative production for the intended pedagogical purposes; c) range (defined as the number of files in which they appeared).

***BE* CADAD Subset: Word Distribution by Frequency and Range**

Context: Billy receives a letter of acceptance from the Royal Ballet School

Characters: Billy, Jackie (Billy’s father), Tony (Billy’s elder brother), Billy’s grandmother

Setting: Billy’s home (kitchen and bedroom)

Range: 38

Tokens: 5908

NB: The following sections may contain spelling or grammar mistakes, as originally written by the learners.

1. General Words

Unit	Frequency	Range
letter	157	38
family	110	33

father	81	35
room	72	34
table	62	33
look	40	24
kitchen	35	23
door	34	22
school	29	23
brother	29	20
bag	22	22
grandmother	19	15
armchair	19	12
cigarette	17	16
pillow	16	15
grandma	15	7
chair	13	12
face	11	8
relatives	11	5
uniform	10	10
expression	10	8
tension	9	6
eyes	9	9
breath	8	6
fingers	8	8

Envelope	7	3
Fear	7	5
anger	2	1
glance	1	1

2. Who

Naming	Frequency	Range
He	293	38
Billy	155	38
His father	65	30
His brother	29	20
The grandma	15	7
The boy	9	6
Tony	2	2
The protagonist	2	2
His dad	1	1
The grandma and the dad	1	1
Jackie	1	1
Billy Elliot	1	1
The kid	1	1

3. What and How

Main Verbs		
Main tenses: Present Simple; Present Continuous; Other tenses: Present Perfect		
look	121	33
sit	51	41
take	35	25
enter	29	18
read	36	37
open	27	25
start	26	16
go	23	21
close	20	19
seem	18	12
wait	18	10
stand	17	16
continue	14	10
drop	12	11
open	12	11
stare	12	8
wear	10	10
hug	9	9
light	9	9
put	9	7

sigh	8	5
bite	8	5
smoke	6	6
sweat	3	2
slam	2	2
tap	2	2
stream	1	1
turn into	1	1
end up	1	1
convince	1	1
glance	1	1
stomp	1	1
ruin	1	1
dry	1	1

Adverbs and Prepositional Phrases (PP) and adverbs related to emotions and feelings

Emotions/feelings

anxiously	21	15
nervously	5	4
tensely	4	4
fearfully	4	3
worriedly	4	2

deeply	2	2
excitedly	1	1
restlessly	1	1
stressfully	1	1
uncertainly	1	1
Manner		
slowly	7	7
heavily	3	2
firmly	2	2
badly	2	1
gently	1	1
continuously	1	1
Temporal		
suddenly	1	1
probably	1	1
Prepositional Phrases		
with and anxious stare	1	1
with his arms folded	1	1
with teary eyes	1	1
with watery eyes	1	1
with a last check	1	1

Adjectives		
Unit	Frequency	Range
mysterious	2	2
yellow	2	2
typical	1	1
Northern	1	1
British	1	1
bated	1	1
palpable	1	1
soft	1	1
high	1	1
odd	1	1
unopened	1	1
unputdownable	1	1
Emotion-related adjectives		
Unit	Frequency	Range
nervous	25	22
worried	25	18
anxious	13	11
sad	9	7
scared	7	7
devastated	4	4
surprised	4	4

agitated	4	2
fearful	4	3
tense	3	3
upset	3	2
shocked	2	2
disheartened	1	1
insecure	1	1
distressed	1	1
perplexed	1	1
hesitant	1	1
anguished	1	1
thoughtful	1	1
preoccupied	1	1
disheartened	1	1
insecure	1	1
sorry	1	1
teary	1	1
frustrated	1	1
despondent	1	1
disgruntled	1	1

Figurative Language

Unit	Frequency	Range
on pins and needles	11	11

tied up in knots	4	4
a nervous wreck	4+6 incorrect use	7
on edge	4	4
to fret / to be fretting	1+3 incorrect	4
jitters (to have the)	2+1 incorrect	3

4. Where

Spatial dimension	Frequency	Range
room	72	33
table	62 Physical description: • <i>Kitchen table</i> (4)	33
kitchen	35 No physical description	22
door	34 Physical description • <i>Sliding door</i> (4); • <i>Yellow sliding door</i> (4) • <i>Glass door</i> (2); • <i>Yellow door</i> (2) • <i>Transparent door</i> (1) Spatial description • <i>The door separating the kitchen from the other room</i> (1)	22
armchair	19 Spatial description, no physical description • <i>Sitting in an armchair next to the bed</i> (1)	12

5. When

Temporal dimension		
Unit	Frequency	Range
while	43	20
after	18	11
before	7	6

6. Culture-specific References (CSRs)

CSR	Frequency	Range
uniform	10 - wearing a school <i>uniform</i> (4); - wearing his <i>uniform</i> (5); - who's still in <i>uniform</i> (1)	8
typical northern British wallpaper	1 Billy walks into the room which has <i>typical Northern British wallpaper</i>	1

The ES CADAD Subset: Word Distribution by Frequency and Range

Context: US TV host Ellen DeGeneres welcomes children in fancy dress during a Halloween-themed episode

Characters: Ellen DeGeneres, three young guests (one girl and two boys), studio audience

Setting: The Ellen DeGeneres Show studio

Range: 29

Tokens: 2766

NB: The following sections may contain spelling or grammar mistakes, as originally written by the learners.

1. General Words

Unit	Frequency	Range
Ellen	126	29

kid	73	18
child	41	12
candies	34	14
costume	27	18
pumpkin	25	19
chair	24	22
box	23	16
audience	20	17
girl	20	11
kids	20	15
dressed	19	11
shaped	19	14
enters	18	9
microphone	18	14
open	17	17
gives	16	12
little	16	10
first	15	14
candy	14	8
puts	13	9
Cecilia	12	8
door	12	11
children	11	8

magician	11	8
slides	11	11
plastic	10	8
sweets	9	7
basket	9	3
Halloween	8	7
bucket	7	3

2. Who

Naming	Frequency	Range
Ellen	126	29
kid	73	18
The child	41	12
He	26	14
She	21	12
The girl	20	11
The kids	20	15
boy	1	1
Cecilia	12	8
Children	11	8
The Presenter	9	3
Clint Eastwood	5	5
They	4	4

Chair girl	1	1
The boy	1	1

3. What and How

Main Verbs Main tenses: Present Simple; Present Continuous; Other tenses: Present Perfect		
Unit	Frequency	Range
enter	23	14
put	21	13
take	17	15
open	17	17
give	16	12
wait	9	9
hand	9	7
approach	8	5
show	8	5
come	8	4
walk	8	7
look	7	7
hold	7	5
appear	7	6
offer	6	6

clap	5	5
stare	4	3
receive	3	3
wear	3	2
insert	2	1
drop	2	1
continue	14	10
drop	12	11
open	12	11
stare	12	8
wear	10	10
put	9	7

Complex verb constructions

Verb+adverb; phrasal verb; verb+direct/resultative adjective

Unit	Frequency	Range
slide open	11	11
walk in	7	6
come in	5	4
step out	3	3
come out	3	2
fill up	3	2
hold out	1	1

join in	1	1
nod shily	1	1
pour into	1	1
cluelessly roam through	1	1
cluelessly stare	1	1
stay still	1	1
step in	1	1
step up	1	1
stomp towards	1	1
stride towards	1	1
walk into	1	1

Adverbs		
Unit	Frequency	Range
cluelessly	2	1
silently	1	1
finally	1	1
Physical appearance based on clothing style		
Unit	Frequency	Range
costume	27 Style and type: • <i>Magician</i> (4), <i>odd</i>, (1); <i>physyque</i> (1), <i>muscled</i> (1)	15

dressed	19	11
	Participial phrases: • .. <i>as a chair</i> (5); • .. <i>as Clint Eastwood</i> (1); • .. <i>like a chair</i> (2); • .. <i>like a literal chair</i> (1); • .. <i>as a seat</i> (1); • .. <i>as a magician</i> (2); • .. <i>as a man</i> (1); • .. <i>for Halloween</i> (1); • .. <i>in a shirt</i> (1); • .. <i>in a costume</i> (1); • .. <i>in all black</i> (1); • .. <i>in large trousers</i> (1); • .. <i>under his jacket</i> (1)	
shirt	8	6
	color: • <i>white shirt</i> (2); type: • T-shirt (1); • <i>Pumped T-shirt</i> (1); • <i>Pumped Long-sleeves T-shirt</i>;	
jacket	8	5
	• muscular jacket (2)	
wig	5	5
	• <i>a grey wig</i> (4); • <i>a white wig</i> (1)	
muscular	5	5
	• <i>muscular jacket</i> (2); • <i>muscular man</i> (1); • <i>muscular physique</i> (1); • <i>muscular torso</i> (1)	

trousers	4 color: • <i>beige trousers</i> (1); size: • <i>large trousers</i> (1); • <i>big trousers</i> (2)	4
pants	2 color: • <i>beige pants</i> (1); size: • <i>comically large pants</i> (1)	2
belt	1 color: • <i>a brown belt</i> (1)	1
tuxedo	1	1
Clint Eastwood -disguised kid	1	1

4. Where

References to spatial dimension	Frequency	Range
studio	8	4
room	8 Static location: • <i>at the side of a room</i> (1); <i>in the middle of the room</i> (1); Motion: • <i>enters the room</i> (2); <i>joins the room</i> (1); <i>walks into the room</i> (1); <i>leave the room</i> (1); <i>roam through the room</i> (1)	5

Locative		
next to	4	3
	<i>A table (1); Ellen (2); the previous child (1)</i>	
behind Ellen	3	3
stand on one side of the set	1	1
placed to the side of the stage	1	1
at the side of a room	1	1
in the middle of the room	1	1
on the stage	1	1
on the right side of the stage	1	1
in their spot	1	1
behind her	1	1
Directional		
enter the studio	8	5
move to the side	3	3
enter the room	2	2
join the room	1	1
leave the room	1	1
go next to Ellen	1	1
roam through the room	1	1

step out of the frame	1	1
step out of the scene	1	1
step out of the view	1	1
stomp towards Ellen	1	1
stride towards Ellen	1	1
Clint Eastwood costume approaches the presenter	1	1
the child approaches	1	1
the two kids get out	1	1
exit the framing	1	1
exit the stage	1	1
to the left	1	1
walk together to the right side (of the stage)	1	1
walk towards Ellen	1	1

5. When

Temporal dimension		
Unit	Frequency	Range
now	8	5
after	5	5
then	2	1

6. CSRs

CSR	Frequency	Range
Pumpkin-shaped	17	16
Magician costume	4	4
Halloween bucket	2	2
Halloween clothes	2	2
Muscles and abs	2	2
Chair-disguised girl	1	1
Chair girl	1	1
Chair of the Obama	1	1
Clint Eastwood costume	1	1
Clint Eastwood-disguised kid	1	1
Comically large pants	1	1
Costume as Clint Eastwood	1	1
Culturist costume	1	1
Empty chair	1	1
Halloween basket*	1	1
<i>Fakes</i> muscles	1	1
Magic Mike costume	1	1
Magician-disguised	1	1
Magician stick	1	1

Pumped long-sleeves shirt	1	1
Pumped T-shirt	1	1
Pumpkin container	1	1
Pumpkin plastic box	1	1
Muscular man	1	1
Muscular physique	1	1
Muscular torso	1	1

*student's misspelling of "basket"

VITS CADAD Subset: Word Distribution by Frequency and Range

Context: Glimpses of daily life in Aagaswadi, a drought-stricken village in the Satara District, India

Characters: Aagaswadi residents

Setting: Aagaswadi village (Satara District, India)

Range: 29

Tokens: 2529

NB: The following sections may contain spelling or grammar mistakes, as originally written by the learners.

1. General Words

Unit	Frequency	Range
two	56	25
People	36	19
Houses	34	19
Desert	24	20
House	23	12

India	22	20
Film	18	17
Street	18	11
Television	18	17
Three	18	15
Institute	17	16
Presents	16	14
Black	15	14
Bovines	14	11
Cart	14	12
Land	14	10
Village	14	10
One	13	10
Pulled	13	12
Screen	13	12
Deserted	12	9
Conditions	11	9
Sitting	11	6
Woman	11	6
Road	10	9
Rocks	10	7
Animals	9	7
Ground	9	6

Man	9	8
Small	9	6
Trees	9	9
Arid	8	6
Bad	8	8
Dry	8	8
Tired	8	5
Area	7	4
Broken	7	6
Made	7	4
Place	7	5
Plants	7	5
Vegetation	7	6
Boys	6	3
Built	6	4
Carrying	6	5
Indian	6	3
Logo	6	6
Oxen	6	4
Roads	6	5
See	6	3
Seem	6	5
Seems	6	4

Stone	6	4
Trained	6	4
Tree	6	6
Wagon	6	5
Water	6	4
Appears	5	4
Background	5	4
Bovins**	5	4
Carriage	5	4
Door	5	5
Empty	5	3
Landscape	5	5
Live	5	4
Looking	5	3
Old	5	5
Scene	5	3
Sits	5	3
Blue	4	4
Carried	4	4
Close	4	2
First	4	3
Full	4	3
Head	4	3

Look	4	4
Men	4	4
Noise	4	3
Poor	4	4
Same	4	4
Shown	4	2
Side	4	4
Stones	4	4
Weak	4	3
Wood	4	4
Abandoned	3	3
Bovine	3	2
Brick	3	3
Bricks	3	2
Building	3	2
Bushes	3	3
Condition	3	3
Desperate	3	3
Family	3	3
heads	3	2
Life	3	2
Part	3	3
Persons	3	2

Rocky	3	2
Sad	3	2
Silence	3	2
Source	3	3
Streets	3	3
Tiles	3	3
Uninhabited	3	3
Vases	3	2
Vew	3	3
Walls	3	3
Wooden	3	3
Writing	3	3
Zone	3	3
Ancient	2	2
Big	2	2
Calm	2	2
Children	2	2
Clothes	2	2
Cows	2	2
Dead	2	2
Depressed	2	2
Devastated	2	2
Devoid	2	2

Dirt	2	2
Distance	2	2
Drought	2	2
Foreground	2	2
Girls	2	1
Goats	2	2
Grass	2	2
Green	2	2
Kids	2	2
Lady	2	2
Metallic	2	2
Narrow	2	2
Porch	2	2
Pots	2	2
Rural	2	1
Silent	2	2
Roofs	2	2
Unpaved	2	2
Passes	2	1
Alley	1	1
Amphora	1	1
Abguished	1	1
Battered	1	1

Branches	1	1
Bulls	1	1
Biome	1	1
Cobblestone	1	1

* Function words have been deliberately omitted;

** Learners' unintentional misspelling

2. Who

Naming	Frequency	Range
A woman	5	5
Three boys	5	5
Some people	4	4
Three people	4	4
Three men	3	3
Two persons	2	1
Two Indian women	2	2
Another woman	2	2
Two old people a woman and a man	1	1
Two people, a man and a woman	1	1
Three kids	1	1
Two kids	1	1
Three indian men	1	1
An Indian	1	1
An Indian man and woman	1	1
Two elderly people	1	1

The girls	1	1
Two girls	1	1
Someone	1	1
There is no one	1	1
A family of six	1	1
A lady	1	1
A few people	1	1
A group of three people	1	1
A group of aged people	1	1

3. What and How

Main tenses: Present Simple; Present Continuous; Past Simple Forms: active and passive		
Single-word verbs	Frequency	Range
cross	2	1
Talk	2	2
Sit	16	9
Look	12	10
Seem	12	9
See	6	3
Do	2	2
Show	1	1
Have	5	4
walk	4	4
Appear	7	6
Enter	2	2
Move	2	2
Carry	6	5

Travel	2	2
Train	6	4
Live	8	7
Pass	4	3
Open	3	2
Tow	3	3
Take	2	2
Change	2	2
Watch	2	1
Work	4	4
Wear	2	2
Reign	1	1
Allow	1	1
March	1	1
Struggle	1	1
Help	1	1
Describe	1	1
Think	1	1
Return	1	1
Stay	1	1
Use	2	1
Multi-word verbs (includes passive forms)		
Pull by	10	10
Carried by	4	4
Pass by	2	2
Look at	2	2

Look like	2	2
Fall off	1	1
Tow by	1	1
Tow along	1	1
Pass through	1	1
Pass by	2	2
Take back	1	1
Hold up	1	1
Roll up	1	1
Roll down	1	1
Heat up	1	1
Fence off	1	1

Descriptive Constructions

Human-related adjective phrases	Frequency	Range
Two old people	2	2
Two elderly people	1	1
Three young guys	1	1
An Indian woman	1	1
Two Indian women	1	1
An indian man	1	1

Three Indian men	1	1
They appear to be very tired and depressed	1	1
They look very tired and weak	1	1
Two people looking thirsty, sad and tired	1	1
More people looking tired	1	1
Some people tired and heated up	1	1
They seem tired and hopeless	1	1
They seem to be devastated	1	1
They seem devastated, anguished and fearful	1	1
A group of aged people	1	1
Disabled people	1	1
The people [...] are desperate	1	1
Topographical, environmental and Setting-based adjective phrases	Frequency	Range
A desert land	5	5
A dry land	1	1
A desert and arid land	1	1
An empty land	1	1
Full of vegetation	1	1
Devoid of [...] water	1	1
A desert road	2	2
A desert zone A	3	3
A desert area	1	1

A desert path	1	1
A desert route	1	1
A desert village	1	1
A deserted area	1	1
A desert landscape	1	1
A deserted land	2	2
A deserted place	1	1
A rural place	1	1
A deserted road	1	1
A deserted, silent and poor village	1	1
A poor village	1	1
An abandoned village	2	2
In bad conditions	2	2
Critical conditions	1	1
In difficult conditions	1	1
In horrible conditions	1	1
In peaceful conditions	1	1
Really bad conditions	1	1
The houses are rudimentary with weak functions	1	1
Weak and dry trees	1	1
An arid desert	1	1
An arid landscape	1	1
The area is very arid	2	2
Semi-arid area	3	3
Bare and sandy landscape	1	1
A dry landscape	1	1
A bare and sandy landscape with a few dead trees	1	1

The roads are rocky	1	1
Dirt roads	1	1
Unpaved roads	1	1
A narrow street full of rocks	1	1
An empty, rocky street	2	2
A narrow street	1	1
A rural street	1	1
A small street	1	1
A tight street	1	1
Isolated desert	1	1
Gardens are bad treated, full of waste	1	1
Very colorful bushes	1	1
Deafening silence	1	1
Abandoned houses	2	2
Houses [...] built with straw or stone and seem abandoned	1	1
Two ancient and broken houses	1	1
Very poor houses	1	1
Ruined houses	1	1
The houses are rudimentary	1	1
Stone houses	3	3
Uninhabited houses	2	2
Two stone brick houses more damaged than the others	1	1
Poorly cared houses	1	1
Poor vegetation	1	1
Few trees, weak and dry	1	1
Very unkempt with dry plants	1	1
Broken pots, withered plants and small trees	1	1

Adverbial constructions related to characters/objects		
Manner	Frequency	Range
Walk straight	2	2
Walking very slowly	1	1
Roofs mostly collapsed	1	1
Struggling daily	4	4
Silently think a lot	1	1
In silence	1	1
Slightly broken	1	1
Temporal	Frequency	Range
While	3	3
Now	1	1
Then	1	1

4. Where

Spatial references	Frequency	Range
On the side	4	4
Outside	3	3
Inside	3	3
Above	3	3
On a step	2	2
Along a desert road	2	2
In the background	2	2
In the middle (of)	2	2

In the distance	2	2
In this place	2	2
On the ground	2	2
On the street	2	2
On a porch	2	2
Along the street	1	1
In the savana*	1	1
In the foreground	1	1
On a room without walls	1	1
In an open structure	1	1
On a carriage	1	1
On big rocks	1	1
In front of their demolished house	1	1
Between the houses	1	1
Returning to their village	1	1
Back to the village	1	1
On the entry	1	1
In front of the entrance	1	1
In what seems to be the road	1	1
In this ghost town	1	1
In an empty rocky street	1	1

*learner's misspelling of "savannah"

4.3 Assessing Priorities in lexical patterns: an ADLAB-informed analysis

Having defined the first quantitative approach to measurable data through the use of a software-assisted method, the focus now shifts to the second paradigm, that is the priorities emerging from the participants' lexical choices in terms of frequency.

On the one hand, the guidelines illustrated by Talaván *et al.* (2024) provided a solid basis for preparing the general pre-task instruction leaflet for the learners with useful tips on what information should be prioritized. On the other hand, the ADLAB principles provided the basis for an in-depth deconstruction of the entire production process in compliance with the major pillars of filmic AD, namely WHO–WHAT, WHERE, and WHEN.

However, this decision is not meant to call into question the credibility of all the existing AD standards; such norms have played a pivotal role in making audiovisual content fully accessible for the blind with the aim of consolidating various trends and practices (Perego, 2017, cf. Remael & Vercauteren, 2007). All these guidelines contain a vast array of best practice recommendations, along with end-user needs or suggestions on style and delivery, as is the case with the United Kingdom (Ofcom), France (*La Charte de l'audiodescription*), Spain (AENOR), and Italy (*SENZA BARRIERE*), only to name a few (Matamala & Orero, 2013). Interestingly, not only have these regulatory descriptive norms sparked scholarly interest in the field of comparative studies, but they have also shed light on the importance of drafting new scientifically proven guidelines (*ibidem*).

In particular, two studies conducted by the Royal National Institute for the Blind (RNIB, Rai *et al.* 2010) and Bittner (2012) have highlighted that, while sharing some obvious similarities, these guidelines may also differ notably in style, whereas a recent contribution by Luque Colmenero (2024) has offered a critical perspective on the ambiguity surrounding creativity and subjectivity, as most guidelines paradoxically encourage subjective techniques, while insisting on objectivity. For this reason, one could plausibly argue that creating AD may be understood either as a rule-based work or as an act of artistic interpretation (cf. Pensabene, 2023), during which the author-describer will have to negotiate between must-have and optional information (Vercauteren, 2021).

As already discussed, the benefits of incorporating DAD into the competency framework for CEFR include developing media literacy, critical thinking skills (Kleege & Wallin, 2015) and understanding the linguistic and communicative process underlying written mediation. It is in this last process that the dynamic interaction between objectivity and subjectivity becomes important, acquiring new analytical significance. During the CADAD task, the former was presented as the long-established strategy for delivering high-quality descriptions, while the latter was strongly encouraged to facilitate the creative writing process, allowing all participants to exercise their full agency in negotiating the complex meanings emerging from the interaction between the visual code and the verbal code (Romero-Muñoz, 2025, *cf.* Chaume, 2004).

After this preliminary discussion, the question now arises as to why the ADLAB guidelines – published under the title *Picture Painted in Words* in 2015 (edited by Aline Remael, Nina Reviers and Gert Vercauteren) – were the preferred choice for the analysis of the scripts.

First of all, the learners' work was never meant for professional use; instead, it was intended to serve a pedagogical purpose based on academic research. For this reason, the ADLAB publication provided a considerable example of “reliable, consistent, research-based guidelines” (Remael *et al.*, 2015), resulting from a three-year research project, meaning that its solid theoretical background would facilitate a deeper understanding of the challenges associated with AD. With reference to both full-length films and documentaries, the ADLAB project placed special emphasis on storytelling, following the principle that audiovisual content should be framed as a narrative-driven experience where stories take shape. In fact, the project aimed to provide a comprehensive experience to make them fully accessible to VIPs, considering the multimodal nature of such stories.

Another interesting aspect is that the ADLAB guidelines, which function more as strategies rather than ready-to-use rules (Vercauteren, 2021:8), further highlight “the idea that audience members are not just passive consumers but actively process and interpret the film or TV series they are watching and thus participate in the creation of its meaning” (*ibidem*). It is therefore necessary for describers to prioritize ST analysis to determine what kind of story they want to build, and to do so, they may consider the three main principles of story-creation (*cf.* Taylor *et al.*, 2015), which can be summarized as follows:

1. Content selection, where authors decide the elements of the story that will include: a) characters, their actions, and experiences; b) temporal and spatial settings that will have an impact on the main plot and the audience's interpretation of it.

2. Narrative structure, that is, when the order of events, the frequency and the duration of scenes are determined; at this point, the characters' main physical/behavioral traits are developed along with spatio-temporal details.

3. Story presentation, based on how the story is visually presented through filming strategies and stylistic choices that will affect the characters' actions or experiences.

These principles indicate that describers should not produce a TT (AD) unless they have analyzed the ST, which is possibly the main strategy surrounding the formulation of the ADLAB guidelines. In other words, once the describers have mastered the ST, they will be able to produce a TT on which audiences can rely to create their mental models of each story, according to their knowledge and background (Remael *et al.*, 2015).

Regarding this, the diagram below illustrates how audiences combine on-screen actions with additional information in the story and the spatio-temporal settings to build up mental models (*ibidem*):

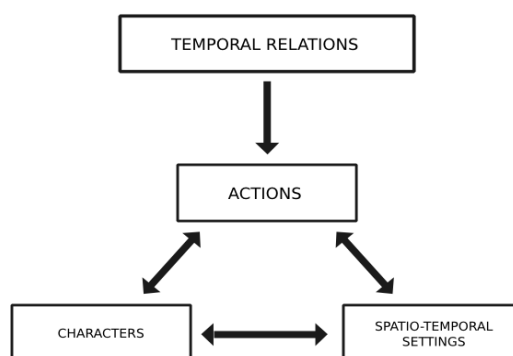


Figure 17: ADLAB guidelines' mental model creation process in story-reconstruction (2015).

The main elements of AD are necessary for audiences to combine their general knowledge with text-specific knowledge (Emmott, 1997), so much so that their mental representations can be updated as the story progresses. This need for coherence is echoed in the ADLAB guidelines, which recommend that describers make decisions aimed at

fostering coherent mental models, even though these will be inevitably influenced by individual interpretation and thus shaped differently throughout the process. Ultimately, when describers go through this complex decision-making process, their narrative choices will still leave a mark on the TT, namely the final AD. In the present study, the story-building process is manifested through and entrusted to creative language and subjective style for didactic purposes.

For many years, a factual approach was regarded as a defining standard of quality in AD, until scholarly research began shifting towards an alternative to objectivity and recognized that no language can be fully defined in objective terms, since communication is strongly influenced by how reality is perceived (Romero-Muñoz, 2023; Vermeulen, 2013), not to mention the fact that sometimes AD has been regarded as solely subjective (cf. Romero-Fresco & Dangerfield, 2022:22) or, at least, lacking convincing explanations of what is truly meant by objectivity (Luque Colmenero, 2024).

There seems to be every reason to believe that treading the fine line between subjective involvement and objective distance is, perhaps, the most critical challenge. When one thinks about it carefully, it becomes clear that the creation of AD depends heavily on the describer's interpretation, meaning that content selection will be influenced by personal interpretation throughout the narrative process, despite all efforts to maintain objectivity. Indeed, this subjective representation can either increase or weaken the impact of AD and is, to a large extent, determined by how language is used for bridging the gap from visual to verbal at both intersemiotic and intermodal levels.

However, the present work does not seek to draw on creative language and style to measure factors such as likability or effectiveness, but to “elicit the students’ awareness that lexical and phraseological competence are of prime importance” (Ibáñez Moreno & Vermeulen, 2013).

If we consider that language played a pivotal role in assessing the final written work, the construction of mental models may have also served a functional purpose for learners acting as describers. Essentially, their guided CADAD activity required them to put their general and text-specific knowledge to the test, which implied gaining an adequate understanding of the visual elements before finding the accurate words to translate their mental representation into linguistic utterances. As Talaván *et al.*, (2024) suggest in their sample guidelines for DAD, describers should always try to visualize the mental images

they intend to convey and, if those images fail to reproduce what they had in mind, they should replace them with other creative solutions.

Interestingly, the participants came up with a wide and remarkable array of creative interpretations of each scene through the use of subjective language, which, as might be expected, was not entirely exempt from mistakes. However, their inaccuracies were marginal and had no real impact on the quality of their productions. The groups demonstrated a reasonable level of understanding of the fundamentals and purposes of AD, with creativity emerging as a complementary dimension of effective communication. Fundamentally, the ADLAB guidelines – which were never shared with the participants – may well exemplify a combination of theory-informed and methodologically rigorous research and were therefore deemed appropriate for the analytical process.

The following sections provide a more detailed analysis of each CADAD subset divided by genre, with a particular focus on their lexical and stylistic choices in relation to the task-specific instructions.

4.4 Analysis of *Billy Elliot* CADAD: enhancing creativity through emotion

The *BE* task was designed to practice emotion-related vocabulary and give the learners full scope for creativity, helping them to develop their language skills while also engaging personally with the material.

As indicated in the first paradigm, the analytical process began with a quantitative observation of general lexical frequency. With reference to the acceptance letter scene, the lists in Section 4.2 revealed that the most frequent lexical unit was the noun “*letter*”, which was commonly preferred over the more neutral “*envelope*”, followed by “*family*”, “*father*”, and “*room*”. Interestingly, “*uniform*” – a key element in CSRs – appeared to be less prevalent both in frequency and range; however, the results were largely consistent with expectations and did not come as a surprise.

4.4.1 Characters, action, language, and style

Character identification was restricted to human figures in relation to their expressed agency within the text, with “*Billy*” obviously appearing as the most frequent hit, ahead of “*his father*” and “*his brother*”. Instead, “*his grandmother*” was used moderately, followed by “*his grandma*” as a hypocoristic form.

Conventionally, describers are requested to decide what relevant physical features may be necessary to present both authentic and unauthentic characters to help the audience create their mental representation of the people/objects in action (Remael *et al.* 2015). In this case, the learners completely absorbed their roles as describers and mentioned the characters as though the audience was already well-acquainted with them. Their descriptions mainly focused on the characters’ psychological traits and the emotionally charged atmosphere, which inevitably diverted their attention away from physical appearance. Some choices to avoid repetition included both neutralizing, strategic glossing – with Billy being mentioned as “*the protagonist*” – and more endearing epithets, such as “*the child*” or “*the kid*”, putting special emphasis on his young age; only a limited number of descriptions identified Billy either as “*Billy Elliot*” (1), or simply as “*Elliot*” (6). Notably, the scene initially revolves around the envelope resting on the table, but the participants decided not to mention that it was closed. Interestingly, while some learners included the letter in their descriptions, a small number failed to notice that it was, in fact, one of the main “actors” within the scene, assuming that the audience had prior knowledge of it.

00:00:00.23 *Billy arrives at home and is anxious to see if the Royal Ballet School accepted him. His family looks at him. Initially his father looks at Billy in the eyes but then he looks away.*

Overall, characters were identified without difficulty with accurate descriptions of their roles and agency within the temporal structure of the scene.

With reference to task-specific instructions on the use of expressive vocabulary, vague verbs were strongly preferred over vivid, descriptive ones (Snyder, 2016) to report on-screen action. The most recurring verbs alluded to perception/cognition – with “*look*” used both as a stative and a dynamic verb – physical action (“*take*”, “*read*”, “*open*”, “*start*”), and movement, expressed by other deictic verbs such as “*enter*” and “*go*”, except for “*stomp*” and “*amble*”, each occurring only once. Therefore, sentences such as “*he goes to*

his room” or *“he puts his bag on the floor”* occurred more frequently than *“he stomps towards a room”* or *“he slams his bag on the ground”*. Likewise, data showed that the Present Simple was the most predominant verb tense as required by AD standards, followed by the Present Continuous (*“Billy’s sitting in an armchair”*), and the Present Perfect Simple (*“he realizes he has been accepted”*).

The learners were not asked to prioritize physical features or general appearance, and description of clothing was limited to Billy wearing a uniform and carrying his school bag. Unsurprisingly, both groups showed great interest in capturing the characters’ psychological dimension. Indeed, tension and hope were the defining features of the scene, which made Billy’s anticipation and eventual relief emotionally powerful. Here, the learners had to describe this emotional crescendo, exercising their ability to translate psychological traits into lexical choices. Results showed an extensive use of emotion-related language with nuanced judgment, although subjectivity seemed less a product of a guided creative process and more a result of their lack of familiarity with AD.

While a limited number of descriptions carried a neutral tone, the majority of them was characterized by emotive language and interpretive extensions, which made them comparatively more impactful and distinctly subjective.

1) 00:00:12.11 *The boy sits down and stares at the letter on the table. He looks at his relatives while biting his lip.*

2) 00:00:15.23 *There is a letter on the table and he knows that it isn't that good at all. He looks at his father in fear. He knows exactly what is going on.*

In the first example, the student described the action to convey Billy’s nervousness; the other excerpt contains supplemental explanation and anticipatory information. Such interpretive strategies – obviously not common in general AD script writing – are clear examples of how lexical choices can influence, or even alter, the audience’s perception. Not only does the phrase *“he knows that the letter isn’t good at all”* convey a highly subjective point of view, but it also presents the audience with information that will later prove to be incorrect. Other interesting examples to define emotional strain were: *“his family is devastated”*, *“with a desperate face”*, and *“it seems that he is shaking”*, to name a few. According to the ADLAB guidelines, identifying how characters and their traits are manifested will help to determine what should be prioritized in the description (Remael et al., 2015:21), which means that in the *BE* scene, those elements are defined by their facial expressions showing tension and anxiety. In this case, the learners’ lexical choices revealed

their natural inclination to name the emotion rather than describing it; therefore, the most recurring choices included emotion-related adjectives such as “*moved*”, “*scared*”, “*nervous*”, “*sad*”, “*upset*”, “*anguished*”, “*anxious*”, and “*worried*”. Nevertheless, less participative approaches could also be observed, in which adjectives were replaced by actual descriptions, e.g., “*He bites his bottom lip nervously*”, “*his grandmother looks at the door with watery eyes*” “*he drums his fingertips on the table*”, also rendered as “*he gently slams his fingers on the table*”.

Following the general stylistic rules of AD, the ADLAB guidelines recommend using “simple noun phrases to pinpoint objects (a military ambulance”, a full shopping basket)”, while also recognizing that variations in sentence structure can make the text more engaging when time permits (Remael *et al.*, 2015). However, the metadiegetic narrative voices occasionally prevailed over the heterodiegetic framework, leading to excessive verbosity or hyperbolic narration, which gave an audio commentary-like quality to the CADAD.

1) 00:01:57.05 *Billy finishes the letter and realizes the big news is turning into something that can ruin his life due to his father's disapproval, namely the acceptance at the Royal Ballet School.*

2) (timing not available) *His brother is sighing above the kitchen sink.*

The sentence “*His brother is sighing above the sink*” appears to be fully fabricated, as Billy’s brother can be seen slowly walking away from the table toward the sink, where he remains standing with his back to the camera, showing no audible signs of distress. However, apart from a limited number of authorial intrusions, another interesting aspect was observing how learners expressed emotions through the use of adverbial and prepositional phrases.

1) 00:00:55.02 *Billy quickly unwraps the letter.*

2) 00:01:07.18 *He starts reading it with a frown on his face.*

3) 00:01:08.15 *He reads with conviction.*

4) 00:00:54.02 *He convinces himself to open it.*

5) 00:01:07.21 *His hands are shaking a little.*

The examples were selected to describe the crucial moment when the Royal Ballet School’s acceptance letter is finally opened. While participants 2 and 5 described Billy’s facial expressions and body language as visible signs of agitation (furrowed brows and trembling hands), student 3 mainly focused on the way the action was performed,

transforming tension into confidence. It should be noted that, although the verb “*unwrap*” is technically incorrect, it suggests that the learners were probably experimenting with vivid vocabulary in their attempt to avoid more common alternatives, such as “*open*”.

Overall, modifiers emerged as lexical priorities in terms of frequency and if, on the one hand, learners showed a certain tendency to name emotions, on the other, they also paid special attention to labelling aspects of material, mental, verbal, and relational processes:

- 1) 00:00:07.22 *His family worriedly stares at him.*
- 2) 00:01:13.16 *Billy sits anxiously at the table with his family.*
- 3) 00:00:44.17 *His brother instead washes the dishes thoughtfully.*
- 4) 00:00:02.05 *Billy enters the house. His father, his grandma and his brother look at him tensely and curiously.*
- 5) *Billy is sitting in an armchair, still wearing his uniform, and opens the letter restlessly.*

The analysis showed that adverbial phrases mainly collocate with lexical units expressing action. In other words, we may say that adverbs were creatively used to describe the characters’ emotional intensity behind their activities and gestures, a strategy also positively encouraged in the ADLAB guidelines, where they appear among the options that can increase text vividness and precision (Remael *et al.*, 2015). However, the aggregate use of adverbial phrases is hardly surprising if we consider that at this stage, secondary school language learners may not have fully developed more nuanced descriptive strategies or may not be familiar with the implications underlying intersemiotic translation.

4.4.2 Spatio-Temporal Dimension

In this short excerpt from the acceptance letter scene, silence can be singled out as a major attention catalyst, sending a thrill of anticipation through the active participants in the scene. In general, no-dialogue scenes allow describers to decide how much AD should be added in the TT creation without having to locate “silent gaps” in the ST (Remael *et al.*, 2015). This does not imply that describers will have to fill in every single pause in the ST or try to communicate all visual aspects; instead, they will be required to operate at an intersemiotic level, taking into consideration any piece of information that is not transmitted

by means of words, music, or other sounds (Hernandez Bartolomé & Mendiluce Cabrera, 2009).

This excerpt was chosen for a run-in CADAD activity because it offered the participants the opportunity to become familiar with AD and practice their writing skills without too many time constraints, as the absence of dialogue would allow them to reformulate their semi-guided creative narration.

As for the reconstruction of the spatio-temporal setting, although the temporal dimension and the spatial component are generally considered as a unified entity from a narratological perspective (Vercauteren, 2012), here time appeared to play a comparatively minor role. The short scene fragment shows a single event occurring at two precise moments, namely, Billy's arrival and the opening of the letter, which takes place in his room seconds later. These moments brilliantly capture the lulls occurring before or after tension through subtle conflict (Truby, 2007), which made it difficult for the learners to insert temporal references, except for a few occasional adverbs or conjunctions, e.g., *while*, *before*, *after* (as shown in Section 4.2).

Conversely, just under half of the CADAD subsets (approximately 46%) contained general references to spatial settings. Overall, the learners made a clear distinction between the two rooms in which the *plan-séquence* takes place, namely the kitchen and Billy's room. Remarkably, one student erroneously identified Billy's room as the living room (*he gets up, looking at his father and goes into the living room*), whereas the remaining 34% merely alluded to Billy moving to an unspecified space where he finally opens the letter (Table 13 below).

Example	Explicit Reference	Example	Implicit Reference
1 00:00:31.17	<i>So, he gets up with the letter still in his hand and he goes to his room.</i>	3 00:00:36.08	<i>He goes into another room and he closes the door.</i>
2 00:00:33.21	<i>He then decides to read the letter in another room, distant from his family's eyes.</i>	4 00:00:31.23	<i>He stands up.</i>
00:00:54.21	<i>Billy then sits on an armchair in what appears to be his room and immediately opens the letter.</i>	00:00:35.08	<i>He is walking to the near room.</i>

Table 13: Implicit and explicit references to spatial dimension.

In the table above, example n.2 shows that the second location was not specified (“*another room*”) until a few seconds later, when Billy had already entered the room. Here, the student correctly followed the spatio-temporal structure of the scene, and this choice proved highly effective in maintaining consistency and continuity even with an inferential formulation (“*what appears to be his room*”).

When a new setting has no narratological-symbolic function, it may simply serve as a background, requiring minimal description in a way that allows audiences to identify it (Remael *et al.*, 2015), and this scene is no exception. However, a limited number of descriptions (13 out of 38) contained general references to home furnishing details, namely the door to Billy’s room and the kitchen wallpaper. The former was described mainly in terms of composition (implicit/explicit, e.g., “*glass door*”, “*transparent door*”) or color/function (e.g., “*yellow door*”, “*yellow sliding door*”) (see list in Section 4.2), whereas the latter was mentioned only once, as “*typical Northern British wallpapers*”.

4.4.3 Figurative Language

Figurative language was singled out as a crucial strategy for developing creative writing skills and promoting authentic engagement with the language. For this reason, the learners were encouraged to incorporate a short list of preselected emotion-related idiomatic metaphors into their CADAD. As illustrated in Chapter 3, these fixed expressions, all drawing from physical states, movement and pressure, were chosen to help them verbalize abstract feelings through cross-domain conceptualizations, while increasing their comprehension and enjoyment (Zabrocka, 2019).

Although the majority of participants had an adequate level of understanding of the suggested metaphorical idioms, the incidence rate of the most frequent expression, namely “*on pins and needles*”, was relatively low, with a percentage slightly below 30% across the total range. However, those who decided to include metaphors did their best to use them appropriately. For example, “*a nervous wreck*” would have ranked as the second most recurrent pattern if not for six documented cases of improper usage, along with “*to be fretting*”, reporting three cases of misuse (“*he looks fretting*”, “*Billy, with fretting*”, “*his hands seem fretting*”).

4.4.4 Film Language

As explained in the ADLAB guidelines, film language can be identified as:

[...] The sum of [...] various film techniques [...] that can be grouped into three broad categories: *mise-en-scène*, cinematography and editing. *Mise-en-scène* [...] includes setting, costume and makeup, and staging. Cinematography deals with how shots are filmed and [...] their photographic qualities, framing and duration. Editing refers to the relations between different shots, which include a graphic, rhythmic, spatial and temporal dimension. In other words, film language determines the form in which the story is told (Perego *et al.* 2015:30).

These film language techniques serve four main functions – namely, denotative, expressive, symbolic and aesthetic – to articulate narratives, elicit emotions, and convey multifaceted meanings to the audience (online). However, describers should avoid making shot transitions explicit and instead simply juxtapose descriptions of different scenarios. Conversely, they can decide to either name the technique (e.g., in slow motion, in black and white, *etc.*) or its function (e.g., in a 30s movie) (*ibidem*) to preserve the scope of the film.

Returning to the *BE* clip, the learners were essentially instructed to avoid direct references to cinematic terms and unnecessary expressions such as “we see”. As for film language, there were only a few instances of references to film-making methods not included in the table, which may have been influenced either by unintentional strategies of defamiliarization or simply by no previous exposure to AD. More specifically, four texts out of 38 emphasized framing with “*move*”, “*change*”, “*shift*” as verb collocates, along with the *N-Gram* aggregate “*the scene changes*”, appearing five times in a range of two files across the entire dataset. The latter occurred only once as “*the scene changes and now Billy is sitting in his room [...]*”. If we consider that “*now*” is commonly used to indicate a scene or shot change in AD, the sentence ultimately ended up being redundant and implicitly repetitive.

Interestingly, the analysis revealed only one case in which cinematic vocabulary was replaced with “*in the other room*”, providing only a small but clear hint at a shot change.

4.5 Analysis of *The Ellen DeGeneres Show* CADAD: creativity through action and physical description

Unlike *BE*, *ES* is structured around a more vibrant and dynamic rhythm. Due to its fast-paced sequences, this excerpt required a different approach, which translated into a more balanced frequency distribution. Admittedly, it was almost impossible for the learners to avoid certain repetitions, especially those connected to character and object identification, to the extent that the most recurring units were “*Ellen*” – the main catalyst behind the course of action – and other common nouns, such as “*kid*”, “*costume*”, “*pumpkin*” and “*chair*”. Conversely, “*Halloween*” appeared only eight times, which does not necessarily imply that a salient lexical feature was disregarded; it is more likely an expression of the participants’ descriptive abilities to avoid redundancy instead.

4.5.1 Characters, action, language, and style

After the first task, the learners transitioned into what could be defined as a more demanding situation. Ellen’s frequent witty utterances and her hilarious exchanges with the children in costume prompted humorous dialogues which considerably reduced natural pauses. This resulted in shorter descriptions and fewer discursive sentences with increased lexical density (e.g., “*Ellen collects some candies and puts them into the girl’s pumpkin-shaped basket*”; “*The kid shows his muscular jacket*”).

It is no surprise that “*Ellen*” appeared 126 times across all 29 files with a 100% range coverage and only 9 times as “*The Presenter*”. Instead, the pronoun “*she*” was used 21 times in total – 15 referring to Ellen and 6 to her guests in costume – reflecting a mixed usage. The girl dressed in the chair costume was the only child to be identified by her given name, which was likely the easiest to remember due to its phonological similarity to Italian and appeared 12 times as “*Cecilia*”, but only once as “*the chair girl*”. Similarly, the boy wearing the Clint Eastwood costume was specifically named “*Clint Eastwood*” five times, contrary to the Magic Mike impersonator. The costumes are two interesting examples of “culture-bound and time-specific elements” (Ranzato, 2014) reflecting two relevant North American cultural phenomena at that time, which combined political satire with cultural parody for entertainment purposes. This means that, even though both Clint Eastwood and

Magic Mike could be recognized within their cultural contexts, the former was comparatively a more familiar figure for Italian learners.

Another important aspect was the verbalization of the on-stage action with a special focus on both simple and complex verb constructions (*i.e.*, phrasal verbs, verb + adverb), as illustrated in Section 4.2. In the first group, the most frequent verbs, such as “*open*”, “*enter*”, “*put*” and “*give*” clearly showed a tendency to use direct language. In contrast, “*slide open*” and “*walk in*” stood out as more complex constructions. Other interesting lexical units or bundles, including verb fragments such as “*stomp towards*”, “*roam through*”, “*pour into*”, and “*peek*” – to name a few – reflected the participants’ effort to convey vividness to a certain extent, although they recorded a near-zero frequency:

- 1) 00:00:26.23 *The kid stomps towards Ellen.*
- 2) 00:01:30.09 *The kids cluelessly roam through the room trying to exit the framing.*
- 3) 00:00:45.24 *Ellen picks up some candy and pours it into the box.*

During the costume parade, as the little guests make their entrance carrying Halloween-themed buckets, Ellen offers them some assorted tasty treats. The most frequent units to describe this moment referred to the bucket, which was mainly identified as “*box*” (23), “*bucket*” (7), and “*basket*” (9). With reference to action, basic verbs were generally prioritized over more vivid expressions, meaning that sentences like “*Ellen gives him some sweets*” occurred more frequently than “*she takes a scoop of candies and puts them into the kid’s bucket*”, where lexical density is traded off with more colorful details.

Additional verbs and expressions used for the candy distribution moment are listed as follows:

- 1) 00:00:44.13 *Ellen tries to reach the child candy box then proceeds to fill it with candy.*
- 2) 00:00:44.22 *Ellen first tries to take the bucket, but instead decides to only collect the sweets and hands them to him.*
- 3) 00:01:21.03 *Ellen inserts the sweets in the box.*
- 4) 00:01:18.18 *Ellen gathers some candies and puts them into the girl's basket.*
- 5) 00:01:18.18 *The presenter also drops some candies in the other kid's bowl.*

Overall, one noticeable feature of the learners’ work for the Ellen episode was a limited use of verb variation, which is not surprising if we consider that, in many cases, this

might be associated with limited contextual exposure, in the sense that the use of bilingual dictionaries may not suffice to integrate verbs naturally.

This can be exemplified in the description of the opening sequence, when the first costume is introduced to the audience. For the majority of the participants, the first child “*enters*” the main stage or simply “*comes out*”, while others focused on expressions like “*comes in*” or “*walks in*”. However, other additional expressions were also “*appears*”, “*makes his appearance*”, “*joins the room*”, or “*approaches the presenter*”, not to mention that in a couple of isolated cases, the participants came up with creative workarounds to more common verbs, such as “*shows up and stomps towards Ellen*”. Only in one case, the sentence “*the door in the background slides open to present the first kid who walks into the scene*” seemed to maintain a slightly smoother flow.

If characters are primarily introduced through their physical features, actions and reactions (Mazur *et al.*, 2015), in certain cases, clothing and accessories can be an important part of the narrative (*ibidem*). Therefore, in this Halloween-themed excerpt, the children’s costumes become the must-have elements of the descriptions, drawing attention to visual appearance as a dominant factor. Both the laughter track and the studio audience’s reactions help to bring out the intended humorous effect, meaning that the learners had to renew their efforts to ensure consistency with onscreen content. In this regard, creativity was manifested through a combination of nouns, adjectives, and prepositional phrases used as triggers for the creation of vivid mental representations to define the target audience’s experience and enjoyment (Holsanova, 2021).

Predictably, “*costume*” was the most frequent noun, followed by “*dressed*”, which served as a common denominator for several low-frequency participial phrases both for the Clint Eastwood-chair couple and for the Magic Mike costume. All three costumes required careful descriptions and provided interesting inspiration for creative vocabulary. Apparently, the first two costumes triggered the learners’ creativity, although the descriptions of the so-called Chair girl exhibited a higher density of exploratory expressions, whereas Clint Eastwood was simply introduced as “*Clint Eastwood costume*” or the “*Clint Eastwood-disguised kid*”.

1) 00:01:04.01 *Another kid comes in and she interprets the chair of Obama.*

2) 00:00:59.04 *Another child then comes out dressed like a literal chair with a cutout on her face.*

3) 00:00:57.21 *A little girl with her head stuck in a chair and a pumpkin-shaped box enters the studio.*

The sentences above refer to the chair costume, illustrating three different descriptive approaches based on subjectivity (1, 2) and neutrality (3); in the first example, the student expressed subjectivity through designation instead of providing a detailed description of the costume, while in the second, the modifier “*literal*” added a personal nuance to what appeared to be more neutral to a moderate degree. Lastly, the third was the only statement that contained a deceptive formulation (*the head stuck in a chair*), despite showing a more disengaged approach.

Conversely, the Magic Mike costume appeared as the “*magician costume*” four times (including the erroneous constructions with a Saxon genitive), and “*magician*” was the highest-ranking unit referring to the costume, with a 28% range coverage.

Overall, the reconstruction of Magic Mike’s athletic build was rendered through adjectives such as “*muscular*” and “*pumped*”. These adjectives were also used to describe the character’s clothes, providing a clear picture of what the costume really looked like: “*muscular jacket*”, “*pumped T-shirt*”, “*pumped long-sleeved T-shirt*”.

Referring to Clint Eastwood, the child was wearing beige underarm-high trousers – reminiscent of the actor’s style – a plain white shirt, and a grey wig. The costume bottoms appeared four times as “*trousers*” and two times as “*pants*”, while only one student accentuated their comedic effect (“*comically large pants*”).

Interestingly, the analysis of this description subset revealed a high concentration of hyphenated word groups used as concise phrases to replace longer constructions, such as “*Clint Eastwood-disguised kid*”, “*pumped long-sleeved T-shirt*”, “*chair-disguised girl*”, “*magician-disguised kid*”, and last, but certainly not least, “*pumpkin-shaped*”, appearing 17 times as the preferred adjectival phrase for the children’s Halloween-themed buckets.

However, if on the one hand, such constructions illustrate how the learners engaged with lexical creativity, on the other hand, a couple of personal interpretations appeared slightly offbeat (“*Ellen pours too much candy*”; “*another kid seems to be the rest of the costume of Clint Eastwood*”), or redundant (“*Ellen is waiting for silence*”, “*the child says what he is disguised*”, “*she’s now having a conversation with the kid*”). Even so, rather than serving as a critique, this analysis aims to provide a fair and balanced assessment of the participants’ work, considering that they were experiencing AD script-writing for the first time; also, the second clip proved somewhat more challenging than the first, due to shorter

pauses and a faster pace, which inevitably called for text condensation. Despite this, their descriptions, far from being disappointing, were well above expectations.

4.5.2 Spatio-Temporal Dimension and Film Language

Notably, only a limited number of spatial references were detected in this scene. Still, the results showed that the terms “*studio*” and “*room*” each appeared eight times in total (with little variation across files). All the other prepositional phrases showed minimal occurrence in the corpus and included both locative states and directional/translative constructions. Examples from the first group include: “*next to Ellen*”, “*behind Ellen*”, and other expressions related to stage blocking¹⁶, such as “*on the right side of the stage*”, “*stand on one side of the set*”, or “*sit at the side of a room*”, to name a few. Instead, “*enters the studio*”, “*move to the side*” were among the preferred configurations to express movement, not to mention a couple of elaborate solutions, like “*striding towards Ellen*”, “*stomps towards Ellen*”, or “*roam through the room*”.

Due to the entertainment-oriented structure of *ES*, where perfect comic timing and dialogues are crucial to maintain its performative rhythm, the only temporal references were “*now*”, “*after*” and “*then*”, each showing a frequency of eight, five and one across the entire range. However, two sentences from the same script – already partially mentioned in Section 4.1 – deserve attention for their peculiar structure containing references to both spatio-temporal settings and film language:

- 1) 00:01:30.09 *The kids cluelessly roam through the room trying to exit the framing.*
- 2) 00:01:35.20 *The kids finally leave the room.*

Although the impact of the message is slightly diminished by a curt nod to technical language (“*exit the framing*”), the expressive tone is retained through adverbs alluding to subjectivity (“*cluelessly*”) and temporal sequences (“*finally*”).

Similarly, three files explicitly mentioning television timing contained the following constructions:

- 1) 00:00:07.18 *Ellen waits for the audience to stop clapping.*

¹⁶ Blocking involves several key elements including the precise locations where actors stand, known as stage positions, as well as their movements, orientations, and interactions (NIPAI, 2024).

2) 00:00:06.15 *Ellen waits for the applause to end.*

3) 00:00:06:00 *Ellen waits for the audience to finish applauding*

Arguably, all these examples describe nearly identical moments that could have been omitted; still, they reveal the participants' *naïve* approach to AD with a clear tendency to fill empty spaces.

4.6 Analysis of *Aagaswadi – Village in the Sky* CADAD

Finally, the learners were encouraged to produce CADAD for a short excerpt from *VITS*, already introduced as an example of “slow cinema”, or more specifically, *cinéma vérité*. The task involved a smaller group of learners (20) and was based on free writing, allowing them to experiment with vocabulary drawing fully from the instructions for *BE* and *ES*, which, in a way, seemed perfectly in line with the film's unscripted action. In this short clip, the absence of dialogue gave the participants more time to elaborate their descriptions, which resulted in a larger amount of recurrent words. As shown in the table, “two”, “people”, “house” and its plural form “houses”, “desert”, “India” and “film” were among the words with nearly full range coverage, along with a considerable number of topographical references, such as “street,” “land”, “village”, “area”. Unsurprisingly, some of the most interesting results at the bottom of the list included other refined choices like “biome”, “amphora” and “cobblestone”, each appearing only once in three separate texts.

Contrary to the first two videos, the *VITS* clip showed the first two minutes of the documentary with opening credits. The latter appeared as a single capitalized/lowercased writing on a black screen – *FILM AND TELEVISION INSTITUTE OF INDIA presents* – providing details on production but not the title. The data analysis revealed that only ten participants decided to include the credits, while the others began their description when the first image of a deserted landscape appeared.

4.6.1 Characters, action, language, and style

In this thematic section, the majority of the outputs exhibited low frequency rates, meaning that the learners used varied terminology to express similar concepts based on

their individual descriptive choices, e.g., “*they appear to be very tired and depressed*”, “*people looking thirsty, sad and tired*”.

A significant example is the adjective “*desert*”, which was one of the highest-ranking units in the general word list despite showing comparatively low results in these concordance lines.

In other words, the frequency rates were highly influenced by a context-based categorization that generated a significant number of isolated instances of the same unit, such as “*a desert land*”, “*a desert path*”, “*a desert route*”, “*a dry landscape*”, “*an arid landscape*”, among others.

All the people appearing in the clip were part of its silent narration and their names were never revealed. Therefore, due to the ambiguity of the scene, the results showed a ubiquitous presence of common nouns referring to unspecified individuals, and all the people who took on-screen action were mainly identified by adjective-noun patterns indicating small groups of individuals, e.g., “*three boys*”, “*three people*”, “*some people*”, “*a woman*”, and “*three men*”. Other expressions with minimal occurrence also included demonyms and adjectival forms referring to nationality (*Indian*), age (*elderly*, *aged*) or indefinite adjectives (*someone*, *some people*).

Verbs were divided into two distinct categories, namely single-word and multi-word verbs, with a high frequency of phrasal verbs. Like in *BE*, the most frequent single-word verbs relied on: a) perception (look, seem, see); b) real action (carry, appear, work); c) movement (walk, pass). However, the contemplative style of the film is also reflected in the use of other verbs expressing the inhabitants’ static nature, such as “*sit*”, which was among the highest-ranking units in the whole corpus.

Conversely, among multi-word verbs, which seemed to capture well the adverse conditions that the villagers are experiencing, the most recurring patterns were the passive forms “*pulled by*” and “*carried by*”, used to describe the ox-drawn cart appearing at the beginning of the film.

Interestingly, the learners found this exercise more demanding than the previous ones, especially when the long silent moments became a realization that their description was the only possible instrument to make the film meaningful to an imaginary blind audience, thus requiring precision.

Notably, while not all of the learners’ suggestions were entirely correct in terms of syntax and grammar, the variety of terms used to describe the first fragment is a clear sign

of their efforts to ensure clarity. Therefore, nouns like “wagon”, “cart”, or even a “carriage” were described as being “*pulled by bovines*”, or “*passing by carrying three people pulled along by oxen*”. This scene is one of the few moments showing active movement, along with the shot of women carrying containers on their heads; otherwise, the film is largely dominated by a sense of weary resignation, with the landscape functioning as the true protagonist.

The learners used a vast array of topographical and environmental adjectival phrases to describe this dramatic scenery, with “desert” appearing as a common denominator: “A desert land”, “A desert and arid land”, “A desert route”, “A desert path”, “A deserted area”, “A deserted, silent and poor village”.

Other constructions seemed to perfectly illustrate either the barren panorama – “semi-arid area”, “the roads are rocky”, or the difficult circumstances that the people are subjected to – “Two people looking thirsty, sad and tired”, “they seem tired and hopeless”, “they seem to be devastated”, “the people are desperate”. In general, the learners succeeded in capturing the gaunt appearance of the villagers, hinting at their discomfort and loneliness (“the animals that seem to be the only ones to provide company are fenced off”), with one exception where the description appeared reductive, if not slightly inappropriate (“a place where disabled people rest”).

Correspondingly, creativity also appeared through interpretive statements with a story-like quality:

- 1) 00:00:25.19 *Inhabitants cannot use means of transport; they use wagons transported by animals.*
- 2) 00:00:09.02 *This is India, a desert path where life can't belong.*
- 3) 00:01:14.18 *People are desperate for something; surely one of the main problems is drought.*

About one-fifth of the participants enriched their descriptions with personal interferences – “they seem desperate due to the water being scarce”, “they carry multiple bowls, probably with some water” – that occasionally took the form of full narratives, as shown in below:

- 1) 00:01:25.13 *People who live in this rural place have a simple life. However, they seem to be devastated. They stay outside to kill some time or they work to help their family, for example raising goats.*

2) 00:00:26.08 *A chariot carried by animals looking like bovines is taking three boys in what seems to be the road to the village.*

The first example goes beyond simple description, both narratively and linguistically, with creativity emerging through extra details about the villagers' everyday lives and the use of the colloquial phrase "*kill some time.*" In the second excerpt – which contains a linguistic inaccuracy (*chariot*) – the student's hedging technique is reflected in the use of "*what seems to be*" and "*they seem to be*", both expressing an epistemic function.

In contrast, a similar proportion showed the concise, nominal style typical of AD syntax. Below are some examples of such sentences from different separate texts, together with alternative versions that conveyed the same concept using different wording and style:

Student A

00:00:06.19 *A dry landscape with sturdy trees.*

Student B

00:00:07.02 *A desert land devoid of any source of water. In the distance there is something that is moving.*

Student C

00:01:33.10 *A brick house with two people and some animals.*

Student D

00:01:35.03 *An ancient house is hosting a large family.*

Student E

00:01:14.01 *Close up of two villagers, sitting in silence and clearly in need.*

Student F

00:01:14.06 *There are two people, one sits on the entry of a house and the other sits on a kind of bench.*

Participants A, C and E chose brief nominal sentences, in which creativity is shown through: a) adjectives, with the landscape and the trees described respectively as "*dry*" and "*sturdy*"; b) an attributive noun (*brick*); c) a film technique (*close up*) and two prepositional phrases ("*in silence*" and "*in need*") to clarify the people's conditions.

On the other hand, participants B, D, and F prioritized a more descriptive style over lexical density; however, phrases like "*there is something that is moving*", "*is hosting*",

“there are two people” and *“a kind of bench”* still made their formulations sound somewhat clunky.

As previously explained, the VITS CADAD dataset revealed that the learners made extensive use of descriptive adjectives referring to people, environment, and buildings. These adjectives were systematically grouped according to their semantic associations into three categories: a) topographical, for environmental descriptors; b) architectural, describing buildings, and c) human-related, for traits related to people (see chart below).

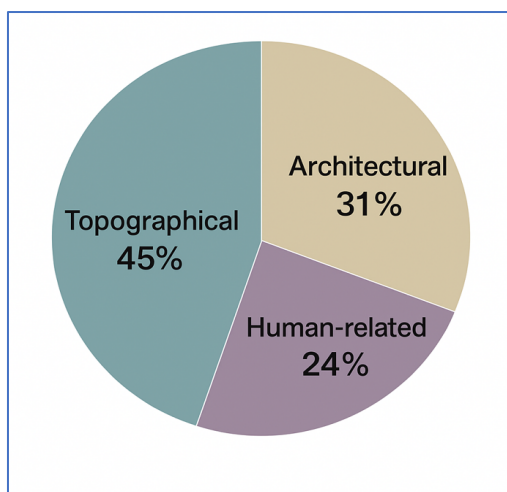


Figure 18: Distribution of adjective use across semantic categories.

As shown in the pie chart above, topographical adjectives account for the highest percentage, which is not surprising due to the landform of the region. This means that the learners decided to either refer to one environmental feature to activate the setting in the user’s mind (Remael & Vercauteren, 2015), or reiterate their landscape descriptions.

By contrast, what is curious about the results is that they included two references to both sanitization (*“the hygienic conditions are at the limit”*) and social status – *“a population of a lower social [sic]”* – which, again, suggests that the descriptive passages were used as a way of channeling their personal thoughts and perceptions.

Finally, adverbs were used very sparingly, with isolated uses of adverbs of manner (*silently, slowly, straight*), degree (*mostly, slightly*), and time (*daily*) occurring across the entire subset, including one adverbial construction (*in silence*).

Likewise, lexical units related to clothing description had a very low frequency, with *“clothes”* attested twice as a general term (*“long colorful clothes”*, *“traditional Indian*

clothes”), as well as “*veil*” (“*veil on their heads*”) and “*dresses*” (“*typical Indian long dresses*”).

4.6.2 Spatial Dimension and Film Language

This final section focuses on how the learners described the spatial dimension of the scene, without considering the temporal domain, which has been deliberately left out due to its limited number of occurrences in the corpus.

Overall, expressions such as “*on the side*”, “*outside*” were certainly more frequent than other specific phrases like “*in front of their demolished house*”, “*on a carriage*”, or “*in this ghost town*”, which appeared only once.

Prepositions like *in*, *on*, and *along* were used extensively to convey both position and relational movement, with phrases like “*returning to their village*” or “*back to the village*”.

Approximately 60% of spatial expressions clearly mention exterior spaces (*outside*, *along a desert road*, *on the street*, *in the distance*) but references to interiors or proper residences were also attested (13%), with constructions like “*inside*”, “*on a room without walls*”, “*in an open structure*”, despite the scene taking place in the open air.

In terms of technical filmic language, about 15% of the subset contained references to camera work or shot types (“*a foreground of an old door*”, “*close-up*”, “*when the image changes*”, “*the scene is shown from three different points of view*”) which, on the one hand, reflected their awareness of how visual elements were incorporated in the scene, but on the other, substantially made their descriptions feel less authentic, as already discussed in Section 4.4.4.

4.7 Frazier® – “Markers” and “Snippets”: Frequency and Application

As already discussed in Chapter 3 (Section 3.12), while creating their CADAD scripts with *VtV* the learners were encouraged to insert both “Snippets” for repeated words and “Markers” as timing cues to flag their entry and exit points, knowing that this would significantly facilitate the editing process.

However, the results of the analysis showed that only 20% of the participants took up the recommendation (Fig. 19).



Figure 19: Marker insertion during the Ellen script writing task on *Frazier®*.

In particular, while the use of “Snippets” remained almost insignificant, a limited number of “Markers” appeared in just two of the CADAD subsets. This minimal use – four cases in *ES* and one in *VITS*, with none in *BE* – suggests that learners may have found this tool either unnecessary or challenging to integrate within their workflow.

1	00:00:00.04	Ellen
2	00:00:28.00	Clint Eastwood
3	00:00:58.17	Cecilia

Table 14: Examples of VtV bookmarks.

Table 14 illustrates how the bookmarks were integrated into the descriptions as bookmarks or personal tags pinpointing the children’s first appearance in their Halloween

costumes. In particular, if four marker reports included the people's names and the timing signaling when their entrance took place, one student chose to mark these moments by adding short phrases instead, such as "*Clint Eastwood comes in*".

Interestingly, after all the submissions had been completed, a short in-class oral survey quickly revealed that 95% of the participants had forgotten to include markers. Arguably, while only a small group decided to maximize their experimentation with the platform, others may have perceived this *bonus* activity as irrelevant to their production, putting in their efforts in the CADAD task instead. Another possible explanation is risk aversion due to low confidence, as some of them may have been unsure about how to add markers properly and avoided using the tool to prevent mistakes. Last, but not least, the participants' engagement may have also been discouraged by possible doubts about the platform design or the teacher's instructions, as learners sometimes tend to be reluctant to ask for clarifications.

4.8 Culture-Specific References

One of the most obvious reasons for learning languages is the ability to communicate with people from different linguistic backgrounds and, most importantly, access their cultural heritage. In a sense, languages are often gateways to other cultures and it is the responsibility of teachers to provide their learners with educational opportunities to gain their intercultural competences besides their linguistic skills, thus promoting multi-culturality (*European Commission, European School Education Platform, 2024*).

Generally speaking, in FLE environments, students learn how to appreciate cultural differences while developing awareness of different cultural perspectives and practices. Specifically, the increasing growth of digital technologies has allowed educators to optimize pedagogical strategies in context-based learning spaces (Psyché *et al.*, 2020) in order to enhance cultural awareness, and it is in these new adaptive situations that DAT-based activities come into play. For example, during the CADAD exercises the learners were inevitably exposed to culture-specific references (CSRs) – also known as *realia* – typical elements of one particular culture which cover a wide range of semantic fields: from

geography to tradition, from institutions and technology to ethnographic and social references (Katan, 2008; cf. Jankowska, 2022: 107).

In particular, although brief and limited in scope, Intralingual CADAD offered the learners considerable advantages to facilitate the development of such competences, providing them with culturally rich inputs (Talaván *et al.* 2024). The short clips included numerous cultural references, some more obvious than others, which the learners may not have consciously recognized or fully taken note of, as it is perfectly normal that certain elements may go unnoticed to trainee describers doing a one-off classroom task. This is mainly due to the fact that learners aged 16-17 are still building both their encyclopedic and conceptual knowledge while also developing their ability to transfer subtle cultural cues from the visual channel to the verbal one. Other possible explanations could be: a) their language level, which may have affected their conceptual knowledge, and b) the extent to which their encyclopedic knowledge – defined by Kecskes (2023) as the broader world knowledge called upon on-demand by situational contexts – influenced their understanding of meaning in this specific situation.

On the other hand, if we consider that intercultural knowledge is gained through direct experience with the affective and behavioral dimensions of other cultures, *i.e.*, understanding why a character's gesture is humorous or significant in a foreign value system (Deardoff, 2006), then it is unlikely for seventeen-year-old learners to have reached a level of intercultural knowledge that allows them to interpret and engage with cultural nuances and not just recognize them. In other words, they may have identified the Clint Eastwood and Chair duo costume as a US cultural reference, but missed the humour behind it before the teacher prompted them.

At the same time, to avoid broad generalizations it is important to consider that learners' level of intercultural knowledge also varies depending on their prior exposure to other languages and cultures, as well as their learning background, not to mention that, while many show curiosity toward foreign cultures they may also be largely influenced by oversimplification and *clichés* (Wang & Ma, 2023).

Now, following these introductory considerations, the three clips offered a broad spectrum of potential (visual) CSRs. Before assessing their presence and frequency in the CADAD sets, they were grouped into type-based categories (Table 15):

<i>BE</i>	<i>ES</i>	<i>VITS</i>
1) School clothes: blazer, tie, shirt marking a typical British school uniform.	1) Halloween objects and sweets	1) Village architecture and layout: Mud and brick houses, narrow pathways, open country yards typical of Maharastrian villages.
2) Household settings: simple furniture and objects marking the regional socio-economic context of a modest working-class Northern English home; the leaf-themed wallpaper; the sliding wooden door with glass panels.	2) Show set: large sliding doors, typical of US talk show programs.	2) Clothing and attire: Maharstrian dress code with sarees and scarves.
	3) Halloween costumes: Clothing and accessories hinting at local cultural/political sensations.	4) Daily rituals: Carrying water in pots.
		5) Animals: Goats, cows.

Table 15: CSRs in the three clips.

The categories are not intended to provide a rigid analytical taxonomy, but rather serve as a general mapping of the most salient CSRs across the three genres, with *VITS* featuring the highest number of references related to traditional objects, clothing, and daily-use objects. However, because the selection is limited to the video excerpts, it should not be taken as a definitive summary of cultural density in the full versions of *BE*, *ES* and *VITS*. Instead, the table functioned mainly as a methodological framework for the quantitative analysis of culturally marked items, rather than as evidence of how they were perceived. Based on the frequency level of recurring lexical constructions, the following discussion will first examine the extent to which these items were identified and rendered in the CADAD scripts for the three genres. Once the CSRs have been mapped, the focus will move briefly to how the learners understood them.

BE: Billy's school attire is perhaps the most explicit cultural indicator, although it was mentioned only ten times and referred to merely as "*uniform*", with no further details

about its specific components. This interpretive choice, with constructions like *“Billy’s wearing a school uniform”*, *“He’s still in a uniform”*, suggests that the learners might have prioritized the emotive dimension of the scene over other descriptive aspects, deliberately omitting details to preserve silent gaps.

Similarly, cues about the setting showed a minimal frequency rate, with very few references to the household environment. Billy’s kitchen has a mid-1980s British working-class style with everyday objects, from utensils to bottles of spiced sauce; these details were omitted by the majority of the learners, except for one of them, who openly mentioned the *“typical Northern British wallpaper”*. Conversely, five participants identified another British furnishing accessory, namely the glass-paneled sliding door, with four descriptions mentioning *“the sliding door”* and only one hinting at a *“glass door”*.

ES: The excerpt alone is a culturally-marked segment celebrating one of the most distinctive Anglo-Saxon traditions, in which the Halloween objects and treats blatantly stood out as unique CSRs to the learners. Both the pumpkin-shaped buckets and Ellen’s sweets showed nearly full coverage as AD denotative elements.

Looking at the show set, the large sliding entry door is arguably a recurrent talk show landmark identified by 10 learners as it *“slides open”*. Instead, the costumes were broken down into clothing pieces and props, with only one description referring to Clint Eastwood’s signature baggy trousers as *“comically large pants”*, followed by *“big trousers”* (2) and *“large pants”* (2). The second costume appeared simply as Magic Mike once, with no explanation or detail provided about its accessories – *“the presenter removes the top of the costume showing a Magic Mike’s costume”* – whereas the *faux* muscular torso replicating the character’s broad chest and prominent biceps were introduced as a *“pumped long-sleeved shirt”* (1), a *“pumped shirt”* (1), a *“muscular jacket”* (2), a *“fake muscular physique”* (1) and a *“muscular torso”* (1).

VITS: Everything in the film is culturally contingent, from the village layout to daily routines, making the clip comparatively richer in culture-bound elements than the others. The learners’ interpretive strategy focused both on the architecture and the rural pathways, with rock and stone-built buildings emerging as the most noticeable and recurring features:

- 1) *“Houses are built with rocks”*;
- 2) *“Houses are built with straw or stone”*;
- 3) *“Houses are made with rocks”*;

- 4) *"Houses are made of yellow brick and stone";*
- 5) *"Stone houses";*
- 6) *"A narrow street full of rocks";*
- 7) *"The roads are rocky";*
- 8) *"The entrance of a house surrounded by many stones".*

If the landscape takes center stage in the scene, the villagers' appearance is of secondary importance, so too are their clothes. The people living in Aagaswadi are shown wearing both random ordinary clothes and cultural garments (*sarees*, long colorful dresses and veils), that were simply described as *"traditional Indian clothes"* (1) or as *"long colorful clothes"*, with only two participants specifying that women wore *"veils on their heads"*. Similarly, the local practice of carrying water containers on the head was captured six times and described as:

- 1) *"A woman holding a pot on her head";*
- 2) *"A lady walks carrying a bin on her head";*
- 3) *"A woman is crossing with a jar above her head";*
- 4) *"Two girls walk straight with vases for water over their heads";*
- 5) *"Two women are carrying multiple bowls, probably with some water in it".*

Lastly, in many rural Indian areas, villages do not only use ox-driven carts, but also milk and feed domestic livestock as part of their everyday life, so it is common for goats and cows to move quite freely around courtyards and, especially, household spaces. This culture-specific aspect was either rendered through general statements or more targeted descriptions, with only one student emphasizing that animals *"are fenced off"*:

- 1) *"Two people and some animals", "some animals outside a house";*
- 1) *"There is a man sitting outside his house, with goats, bovines and chickens next to him";*

Overall, despite the last clip prompted the widest range of context-related cultural references, all the three CADAD sets revealed that the learners' approach remained unchanged across the genres, with subtler details often omitted.

It is worth remembering that, before watching the clips, the learners had been informed about CSRs and were interviewed in general terms (see Chapter 3, Section 3.10). Their responses showed that they were intuitively aware of the concept, although they could not always articulate it in their descriptions. Their perceptions mainly framed CSRs as: a) stereotypical symbols (around 55%), such as the *Eiffel* Tower for France and Central Park for New York – that is, in fact, a recurrent filming location – and b) visible things (around 45%), like traditional clothing, food, or monuments. During a post-task discussion, the majority of them were able to recognize the most obvious culture markers for each clip, mentioning Billy's school uniform, Halloween objects and the water-carrying ritual as explicit references. By contrast, while they could easily identify Clint Eastwood as an A-list celebrity, it was necessary to draw their attention to the humorous meaning of the chair, as the episode of Eastwood talking to the empty chair dates back to 2012 and was not as culturally resonant in Italy.

However, this guided observation can potentially help learners to see culture in action and encourage their critical-thinking, as the exercise is indeed another interesting opportunity to express creativity while analyzing, interpreting and translating meanings into words.

4.9 Framing the emerging priorities in the CADAD scripts

This corpus-based analysis sheds light on the predominant lexical and stylistic patterns across the three CADAD subsets. Upon examining the main lexical and stylistic constructions across the three genres, this section will briefly discuss the emerging lexical priorities drawing on *AntConc* frequency counts and concordance data.

Starting from *BE*, all the descriptions bring psychological characterization into sharp focus, which is substantiated by the learners' repeated use of emotion-related adjectives, such as *scared*, *nervous*, *sad*, *upset*, *anguished*, *anxious* and *worried*.

At the level of verb choices, there is a clear preference for general actions for expressing human agency, meaning that the learners favor vague verbs over vivid ones. On the other hand, modifiers, such as adverbs and prepositional phrases emerge as the main creative effort to label emotions (*desperately*, *anxiously*, *nervously*), qualify actions (*gently slams his fingers*) and add interpretive nuances (*with watery eyes*). This is particularly

significant in terms of stylistic consistency, as it demonstrates how the paucity of colorful verbs is compensated by the creative use of verb modifiers that make descriptions highly subjective. Arguably, evaluative comments (such as “*he knows the letter isn’t good at all*”) and embellished details become a narrative leitmotif, echoing the kind of patronizing retelling discussed by Bardini (2022). However, these results are perfectly acceptable, considering the learners’ limited exposure to AD, and the exploratory, non-professional character of CADAD, in which visual content is more often interpreted than simply described.

Conversely, in *ES*, the learners’ style is slightly more objective, although action and physical descriptions are not given equal attention, with a preference for simplified phrasal verbs. Conversely, once again, stylistic elaboration relies on noun modifiers, so that basic verbs are largely counterbalanced by adjectives emphasizing physical appearance. This is also true for objects and costumes, often referenced through hyphenated adjectival constructions (Clint Eastwood-disguised kid, chair-disguised girl, and pumpkin-shaped bucket). Indeed, the use of condensed noun phrases is an effective strategy for handling the fast-moving scene, demonstrating the learners’ effort to summon their creative and critical-thinking skills for genre-specific descriptions.

Finally, in *VITS*, where the landscape is prominent compared to individual characters, the learners’ creativity mainly stems from an extensive array of topographical modifiers as noun adjuncts (a brick house, a stone house) or adjectives (desert path, dry area), followed by adjectival constructions describing the people’s physical and psychological states (tired, hopeless, desperate). By contrast, attention to verbs and verb modifiers is relatively limited.

Following this discussion, the most recurrent noun and verb modifiers are hereafter grouped into categories, with percentages indicating their frequency (Tables 16 and 17):

Clip	Physical adjectives %	Psychological/emotional adjectives %	Hyphenated %	Topographical/Environmental %
Billy Elliot	12.5	72.5	2.5	5
The Ellen DeGeneres Show	52.5	12.5	27.5	7.5
Aagaswadi	10	27.5	2.5	42.5

Clip	Verb modifiers %
Billy Elliot	70
The Ellen DeGeneres Show	20
Aagaswadi	10

Tables 16 and 17: Percentage breakdown of the most recurring word types in the three CADAD sets.

Overall, the learners tend to adapt their language creatively to different genres when their writing choices are guided by specific requirements and their text is structured in context (Talaván *et al.*, 2024). This structured approach allows them to identify linguistic genres more easily, while also fostering creative writing (*ibidem*).

What teachers could do to guide learners' natural inclination is implement Task-based Language Teaching (TBLT) activities with targeted instructions in modifiers use, for example, where learners can manipulate, produce or interact in their L2, while their attention is focused on mobilizing their grammar, and especially their vocabulary to express meaning (Nunan, 2006). Indeed, the widespread integration of such activities in DAT learning environments can facilitate the development of descriptive writing skills in multimodal contexts.

4.10 Evaluation process: learners' language proficiency through AD Scripts

As previously noted, the AD test was subject to time constraints, spanning only two days for each group. Therefore, the aim is not to measure the longitudinal development of the learners' language level, but rather to assess their writing skills in L2 as evidenced within this specific context.

Central to this discussion is the assessment of integrated skills, as defined by the CEFR/CV (2020), with special emphasis on creative writing as a type of mediation. In this context, AD is also perceived as a mediation strategy for learners (Navarrete, 2024) and, possibly, a hybrid space at the intersection of two categories: "Mediating a Text" and

“Creative Writing Production”. For both categories, the assessment is based on the “can-do” descriptors associated with B2 and C1 proficiency levels and adjusted for the purpose of this study. The following tables illustrate an adaptation of these competences to AD scriptwriting (L2<L2) and do not provide a full account of all descriptors within the CEFR hierarchy, much less the indicators of other types of mediation and writing tasks.

CEFR-informed descriptors adapted to AD scriptwriting (B2–C1)

Mediating a Text (CEFR/CV 2020)

B2	C1
Can convey and summarize the relevant points in complex and source content	Can convey relevant elements, highlighting those requiring further attention
Can describe significant information from structured sources	Can summarize and restructure complex content
Focuses on clear and accurate transfer of meaning	Identifies implicit meanings, integrates evaluation

Creative Writing Production (CEFR/CV 2020)

B2	C1
Can give clear, detailed descriptions on a variety of subjects related to their field of interest	Can produce clear, detailed, descriptions with refined vocabulary
	Can incorporate idiom and humor, where appropriate

Tables 18 and 19: Key differences between B2 and C1 proficiency levels in mediating a text and creative writing production (adapted from CEFR/CV, 2020).

These descriptors illustrate the extent to which learners select, organize, and creatively reformulate relevant information. Crucially, the 2020 CEFR/CV update redefines mediation not just as a cross-linguistic transfer; conversely, it is a cognitive process of "languaging" at the heart of the action-oriented approach, where learners act as "social agents" using language as a primary vehicle for communication (CEFR/CV, 2020). Such an approach is particularly effective in collaborative settings, where learners must "share different inputs, explain their information and work together to achieve a goal" (CEFR/CV, 2020:36).

Furthermore, the task integrates creative writing, defined as "personal, imaginative expression" that utilizes language to "arouse readers' visualization" (Barbot *et al.*, 2012). Consequently, learners are not simply reproducing information but exercising critical-

thinking and creative agency and their scripts are the result of a deliberate decision-making process regarding content selection and discourse organization.

A close examination of the corpus suggests their ability to use varied vocabulary and construct complex utterances. A limited number of marginal patterns of difficulty are localized both in the choice of verbs, which occasionally lack vividness, and very few misuses of suggested figurative language prompts. Despite these minor inconsistencies, the results are highly reassuring. The learners demonstrate strong linguistic competence and a significant capacity to elaborate on visual input, resulting in descriptions broadly consistent with the B2 level and, in several instances, approaching C1.

Chapter 5 – Qualitative Research Methodology

Over the last decades, AD has transitioned from a mere accessibility tool to a prominent subject of academic inquiry. More specifically, there has been a recent surge in scholarly attention that has prompted researchers to investigate its usage through a multitude of theoretical frameworks, moving beyond technical standards to explore its deeper discursive layers. A number of qualitative studies have postulated a convergence between cognitive narratology and the construction of mental models through the information included in AD (Fresno, 2017; Vercauteren, 2012; Geerinck & Vercauteren 2020; Vercauteren, 2021); others have approached AD from a functionalist perspective (Mazur, 2020, 2024; Marra, 2022), building upon the *Skopos* rule to direct content selection toward a cohesive narrative structure. Correspondingly, Braun (2007) has conceptualised AD as a discourse-based, inferential form of multimodal mediation, whereas other scholars have analyzed how different languages negotiate “televisual dynamism” (Bourne & Hurtado, 2007), as well as the rendition of culture-bound elements in translated AD (Mazur, 2014; Jankowska and Szarkowska, 2015; Rizzo & Spinzi, 2023; Rizzo, 2025). Along similar lines, while much of the existing literature focuses on translation strategies for descriptive analysis (Bardini, 2020, cf. Molina and Hurtado, 2002; Matamala and Rami, 2009; Mazur, 2014, Sanz-Moreno, 2017, 2018), a specialized classification of AD strategies was notably absent until Bardini (*ibidem*) provided a taxonomy of AD techniques (ADT), based on Molina and Hurtado’s (2002) functional approach to translation techniques.

Extending this functionalist perspective further, a distinctive body of literature has taken up the tenets of Systemic Functional Grammar (SFG) to deconstruct the meaning-making processes within TS (Matthiessen, 2014a, 2014b; Wang, 2021) and AVT (Taylor, 2003, 2016). Its application to AD has been relatively overlooked to date (Piety, 2004), but its principles are directly pertinent to this study, orienting the analysis towards an evaluative model of AD. Against this backdrop, and building upon the pedagogical importance of CADAD, this chapter introduces a complementary methodology for the qualitative analysis of the learners’ scripts beyond the classroom context. If Chapter 3 provided a top-down assessment through the teacher’s look, in this chapter I take an analyst’s approach to breaking down the descriptions. The discussion is no longer concerned with didactic evaluation, but with how the learners themselves approached the text and what they chose to value within it. First, I will examine a number of representative samples from each genre-

specific corpus to assess the learners' intersemiotic competence in rendering the most salient visual signifiers. These renderings are then interpreted along a creativity continuum and classified according to Bardini's ADTs and Rizzo and Spinzi' (2023) creative strategies, before assessing their functional adequacy.

Finally, the *ATTITUDE* system of the Appraisal Theory (Martin & White, 2005) is layered in to account for evaluative and subjective stances in the learners' language.

Within the framework of the present study, the application of the *ATTITUDE* subsystem of Martin and White's (2005) Appraisal Theory aims to investigate how learners construct evaluative meaning in their AD scripts by deconstructing their language according to well-defined semantic categories. Since the primary purpose of AD is to transform visual information into meaningful verbal representations for B/VIP audiences, the analysis seeks to determine whether the learners' linguistic choices successfully reconstruct the semantic and experiential content of the visual scenes while remaining functionally appropriate to the communicative purpose of the genre. Particular attention is devoted to the three semantic domains of *AFFECT*, *JUDGEMENT* and *APPRECIATION*, which provide a systematic framework for examining how emotional responses, human behavior and perceptual qualities are encoded linguistically. Through this analysis, the study explores the extent to which the learners rely on evaluative language to verbalize visual meaning, negotiate subjectivity, as well as to mediate multimodal experience for audiences who cannot directly access the visual channel.

In brief, the objective is not simply to assess the learners' linguistic competence, but to reveal how their evaluative choices contribute to the effective reconstruction of the experiential and interpersonal dimensions of the audiovisual text, thereby demonstrating the close relationship between creativity, functional adequacy and accessible communication.

5.1 Analyzing Functional Adequacy and Creative Modulation of AD Scripts

Notably, a large number of researchers agree that translation and AD are closely related to each other (Braun, 2008), and, although the cross-modal nature of AD sets it apart from other forms of translation (*ibidem*), we may assume that it is, in fact, a type of

translation of verbal signs. Interlingual translators generally follow global methods that “boil down to translation techniques [...] in the linear process of target text production” (Harmon, 2019:9); a hallmark of this process is the ability to generate more than one workable TT and the judgement to isolate a definitive version (Pym, 2003).

These approaches have been the subject of various, and often contested, taxonomies (see Baker, 1992) which scholars frequently use for comparative analysis. In a broader sense, just as translators can produce multiple viable translations of the same ST, describers sift through various AD renditions of the same AVST before deciding on the final version.

However, a dedicated taxonomy for AD did not exist until Bardini (2020) extended Molina and Hurtado’s (2002) techniques – a functional integration of earlier scholarly approaches – and readapted them into fourteen Audio Description Techniques (ADTs) that prioritize the intersemiotic dimension over cross-linguistic transfer. Such techniques are the result of strategic choices and their validity is contingent upon various parameters ranging from context to purpose (*ibidem*). Following this rationale, this model serves a dual purpose: first, it positions AD as a descriptive act in its own right; second, it leaves considerable room for interpretive modulations.

One of the advantages for such an approach is that interpretive strategies – though often discouraged by official guidelines – reduce the inference load on the part of the end user and are actually better received (Bardini, 2020; cf. Walczak, 2017; Vercauteren & Orero, 2013)

This concept resonates with the notion of creativity in my didactic proposal, and is what drove the need for a second, complementary typology modeled after Rizzo and Spinzi’s (2023) approach – adapted from Schjoldager *et al.* (2008) and Holst (2010) – to grading AD strategies by their creative potential. Although originally tailored for culture-bound markers, their taxonomy lends itself well in this context, covering the full range of descriptive details within the scripts. A personal adaptation of the fourteen ADTs is presented below, organized by levels of creativity rather than following the original sequence proposed by Bardini (2020: 279).

Audio Description Techniques (ADT) adapted from Bardini (2020) and Rizzo and Spinzi (2023).

Non-creative	1. Reduction (brief or inexistent AD)
	2. Generalisation (superficial description)
	3. Particularisation (use of precise terms for specific elements)
	4. Iconic description (denotative AD)
	5. Technical description (use of technical terms, e.g., cinematic language)
Creative	1. Compensation (asynchronous description)
	2. Integral iconic description (replace naming with description)
	3. Modulation (change of focus in AD)
Highly creative	1. Amplification (introduction of extra details)
	2. Substitution (interpretation of extra-linguistic elements)
	3. Adaptation (substitution of CSRs from AVST culture with one from AVTT culture)
	4. Cognitive Description (interpretation of an image of technique)
	5. Creative description (interpretive AD at a connotative level)
	6. Variation (deviation of standard language)

By way of clarification, it is important to distinguish this analysis from the language assessment covered in Chapter 4 (Section 4.9). Despite the didactic nature of AD, the scripts here are not evaluated on the basis of vocabulary acquisition in written production; instead, they are examined in relation to how effectively learners apply AD strategies to preserve function and express meaning creatively. This analysis singles out the highlights of each clip and, specifically, the core elements that need to be explicitly described rather than simply “inferred”, whereas the creativity scale is influenced by the specific demands of the three genres. Following a micro-to-macro approach, I first introduce five representative segments from each genre-specific corpus to show how the learners handled the most important visual signifiers. These, in turn, correspond to multimodal domains that are semiotically dense and central to the construction of intermodal meaning: the letter and the emotive dimension in *BE*, appearance, action and intercultural references in *ES*, and slowness and landscape in *VITS*. Accordingly, the selected segments are associated with ADTs along a creativity continuum – ranging from low to high – and evaluated based on their functional adequacy. For the latter process, I take into account whether the descriptive elements: a) are not only inferable, but also substantiated by visual evidence (Starr & Braun, 2021); b) enhance understanding, and c) serve the communicative purpose (*Skopos*) of the scene. Broadly speaking, while we may condone or take issue with the describer’s choices, we must remain mindful of the fact that certain details are the *sine qua non* for the construction of meaning in AD and cannot be omitted.

Focal elements: tension and letter

- 1) [00:00:15.02] *He's a nervous wreck and, one by one, he looks at his family, while he gently slams his fingers on the table.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Cognitive description	High	✓ Yes	✓ Preserves expressive function; communicates tension	The idiom “a nervous wreck” is a pre-set prompt, here used as a strategy to capture Billy’s inner tension, which translates into his tapping on the table

- 2) [00:00:54.15] *He opens the paper and, before reading the letter, he sighs and prays for a second.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Amplification	High	⚠ Partial	✓ Aligns with Skopos, though partly based on non-inferable amplification	Over-interprets emotional state

- 3) [00:00:21.07] *The air is very tense and everyone looks at him straight in the eyes, curious to know.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Creative Description; Amplification	Medium to High	⚠ Partial	✓ Aligns with Skopos; builds narrative tension	Over-interprets intention

- 4) [00:00:04.19] *He is nervous because there is the letter from the Royal Ballet school with the results of his audition and his family look at him badly because they don't agree with his dream.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Substitution; Amplification	High	⚠ Partial	⚠ Partially fulfils the Skopos; anticipates narrative information and disrupts tension	⚠ Partially aligns with Skopos; over-explicit rendering that introduces non-visible information and reduces narrative tension

- 5) [00:00:00:10] *The protagonist enters the kitchen after school and he is surprised by what is on the table: a letter. His family look at him and seem anxious about what the letter might contain. Billy approaches the table; he takes off his backpack and sits down. He looks anxiously at the letter but doesn't open it.*

ADT	Creativity level	Visual evidence	Functional adequacy (skopos)	Comment
Iconic Description	Low	✓ Yes	✓ Fully fulfils the <i>skopos</i> ; maintains narrative tension	Clear phrasing; minor interpretive elements ("surprised", "anxious") remain plausible and do not disrupt tension

ES

Focal elements: physical appearance, action and culture-bound elements

- 6) [00:00:24.07] *As the door opens, a cute little child dressed in a shirt and comically large pants with an empty candy box comes in walking towards Ellen.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Creative Description	Medium	✓ Yes	✓ fulfils the <i>Skopos</i> ; maintains performative tone and visual incongruity though not complete	Effectively conveys humor through exaggeration, but omits key visual details (e.g., wig)

- 7) [00:00:57.21] *Another child then comes out dressed like a literal chair with a cutout on her face.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Iconic Description; Particularisation	Low to Medium	✓ Yes	⚠ Partially fulfils the <i>Skopos</i> ; the humorous effect is underdeveloped	Accurate description of the costume, but the humorous reference is only partially conveyed

- 8) [00:01:04.0] *Another kid comes in and she is dressed as the chair of Obama.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Substitution	High	⚠ Partial	⚠ Partially fulfils the <i>Skopos</i> ; prematurely reveals the reference, reducing humor	Correct identification of the cultural reference, but premature delivery disrupts the punchline

- 9) [00:01:41.04] *Another child dressed as a magician comes in. Ellen rips off the little child magician costume, revealing a fake muscular physique.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Iconic Description; Particularisation	Medium	✓ Yes	✓ Fulfils the <i>Skopos</i> ; preserves the humorous reveal	Accurately renders the transformation and preserves the humorous reveal without premature interpretation

- 10) [00:01:00.18] *Another child comes onto the stage carrying a large chair, her face peeking through the backrest.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Creative Description; Particularisation	Medium to high	⚠ Partial	⚠ Partially fulfils the <i>Skopos</i> ; prematurely reveals the reference, reducing humor	Well-balanced. Creativity enhances humor without over-interpretation

VITS

Focal elements: slowness and landscape

- 11) [00:00:48.13] *Close-up of a battered house in the village; the roads are rocky and everything looks ruined.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Amplification; Technical description	Medium to high	✓ Yes	✓ Fulfils the <i>Skopos</i> ; conveys the harsh environment	Combines technical framing with evaluative language; slight over-generalization

- 12) [00:00:53.16] *The roofs, mostly collapsed or missing, allow a glimpse of the interiors of the houses.*

ADT	Creativity level	Visual evidence	Functional adequacy (Skopos)	Comment
Particularisation; Iconic Description	Medium	✓ Yes	✓ Fulfils the <i>Skopos</i> ; conveys the living conditions	Communicates structural decay; no over-interpretation

13) [00:00:09.02] *This is India, a desert path where life can't belong.*

ADT	Creativity level	Visual evidence	Functional adequacy (<i>Skopos</i>)	Comment
Creative Description; Amplification	High	✗ Not fully grounded	⚠ Partially fulfils the <i>Skopos</i> ; expressive, but not quite accurate	Expresses evaluation at a connotative level; goes beyond what is visible

14) [00:00:09.15] *A desert land, full of vegetation, where peace reigns, where the silent deafening and the man does not disturb.*

ADT	Creativity level	Visual evidence	Functional adequacy (<i>Skopos</i>)	Comment
Creative Description; Modulation	High	⚠ Partial	⚠ Partially fulfils the <i>Skopos</i> ; lack of details to support the description	Highly poetic and modulated, but introduces non-inferable elements

15) [00:01:25.13] *They stay outside to kill some time.*

ADT	Creativity level	Visual evidence	Functional adequacy (<i>Skopos</i>)	Comment
Amplification; Variation	High	⚠ Partial	⚠ Fulfils the <i>Skopos</i> ; introduces an inferred intention	Introduces inferred purpose through a colloquial expression

This tabulated deconstruction is limited to the representative segments and therefore does not claim to offer an exhaustive account of the entire corpus of the learners' scripts. Instead, these excerpts serve as illustrative instances of broader tendencies within the three genre-specific sub-corpora. Methodologically, the aim is not to generalize from isolated examples, but to show how recurrent descriptive patterns can be interpreted through a creativity gradient and the extent to which the learners' choices align with the functional requirements of each scene.

5.2 Overall Discussion

Looking at these segments as a whole, it is clear that the learners leaned into the methods proposed in Chapter 3, even if their execution remained inconsistent. In all the

three genres, the segments reflect an inclination to favor interpretive solutions over denotative descriptions, as evidenced by the recurring use of techniques like *amplification* and *creative description*. That said, the task was a constant tug-of-war for the learners, who tried to negotiate the fine line between creative expression and functional adequacy, as defined by the framework in Chapter 3.

Starting with the narrative-expressive functions in *BE*, the proposed methods suggested preserving both the scene's slow burn and the impact of visible emotional cues. The selected examples suggest that the participants picked up on the scene's emotional charge but – at the same time – they felt the need to over-explain and spell out motivations that were already inferable. For this reason, phrases such as “*they don't agree with his dream*” border on a quasi-authorial voice that compromises the *Skopos* of sustaining tension and delaying the reveal, although this cannot be generalized to the entire sub-corpus.

In *ES*, where priority was to be given to vivid action verbs, culture-bound references and visual incongruity, the selected examples show mixed results. Some learners acted on the suggestions to highlight specific irony triggers, especially in descriptions that zoom in on costume details without overtly naming the reference too early. At the same time, the examples also reveal two deviations from the suggested method: one was too neutral to let the humor emerge (example 6), whereas in example 8, a premature substitution demonstrated how an early reveal of cultural markers may rob the gag of its payoff. In other words, while the humorous intent was successfully recognized, there were occasional difficulties in fine-tuning the delivery; consequently, the entertaining purpose is partially fulfilled.

The *VITS* clip reflects the informative and expressive functions of the scene, encouraging learners to move away from neutral phrasing and to experiment with light figurative language, lexical variety, and even mildly poetic solutions, while still remaining faithful to the scene's observational nature. Here, the analyzed segments show various solutions ranging from highly precise, observational renderings (example 12) to more interpretive formulations, as in 13 and 14, echoing the so-called “synesthetic style” (Luque Colmenero, 2024); this suggests that the learners took full advantage of the creative freedom that the clip afforded them. However, the examples also flag up that such freedom sometimes leads to excessive abstraction or ungrounded evaluation that may weaken referential precision, which means that the *Skopos* is also partially fulfilled in this case.

Overall, learners are able to identify the challenges of each genre, but their responses vary considerably in terms of control. Interestingly, they appear more comfortable layering meaning onto a scene rather than reducing it, which results in emotional over-explication in *BE*, clunky humor in the *ES*, and poetic modulations in *VITS*. At the same time, the data also include several examples where creativity remains visually justified, functionally appropriate, and genre-sensitive. These cases are especially important because they show how the learners negotiate their choices between *implicatures* and *explicatures*, that is assumptions that they want to render more or less explicitly (Sperber & Wilson, 1995; cf. Braun, 2007), in relation to visual signifiers.

Ultimately, it appears that the didactic instructions were clearly taken up, even though their functional implementation proved slightly more challenging. It is here that the analytical framework becomes useful: by blending ADTs with a creativity gradient and a functional evaluation, we can see that creativity itself is more than an “all-or-nothing” choice. It is not simply present or absent, but it can be more or less effective when it is visually earned. In this respect, CADAD can successfully encourage creative writing, although we should remember that the true lesson lies in reconciling such creativity with the communicative purpose.

5.3 The Architecture of Subjectivity: Appraisal Theory

The focus now turns to the Appraisal/ framework. Far from being a detachment from the analytical process, this stage represents the methodological culmination of the study, where creativity comes full circle within its didactic goals. The Appraisal Theory offers a nuanced perspective on how learners cope with the emotional and evaluative load of the clips through their utterances. As an extension to the lexico-grammar wording scheme of M.A.K. Halliday’s Systemic Functional Grammar (SFG) model (1994), the Appraisal Theory was first advanced as a theory for language evaluation based on its intended functions in social contexts (Pensabene, 2023; cf. Martin, White *et al.*, 1993, 2005) to determine how speakers pass judgement and negotiate evaluations with their addressees (P.R.R. White, 2022; Thompson, 2014). Here, this framework is largely extended to describers and converges directly with the didactic purposes of the study. Specifically, it signals a level of English proficiency at the B2/C1 threshold, which calls for nuanced, evaluative language

use – a task that demands considerable linguistic maturity. This maturity shows up most clearly in the learners' creative agency, where their lexical and stylistic choices imbue the entire AD process with a sense of subjectivity. As already discussed in Chapter 4 (Section 4.2), defining what constitutes a subjective judgment versus an objective interpretation is notoriously difficult (Mazur & Chmiel, 2012). Udo and Fels (2009) contend that total objectivity is impossible, while Hyks (2005) goes further by noting that AD is always a subjective act and describers will always interpret visual information in distinct ways (cf. Luque Colmenero, 2024). As already discussed (Chapter 2, Section 2.8), creativity is also an intricate concept to pin down in AD, ranging from interpretive, artistic solutions to subjective renditions (Romero-Muñoz, 2025); however, in this study, it is regarded mainly as an expression of subjectivity resulting from linguistic choices. Following the ATTITUDE system of Appraisal, such choices are examined in terms of AFFECT (expression of emotion), APPRECIATION (evaluation of things), and JUDGEMENT (evaluation of behavior) which, in turn, help pinpoint the subjectivity-objectivity gradients based on their affective responses, cognitive evaluation and behavioral intentions. However, a clear delineation of its theoretical framework is a necessary first step to fully embrace its practical application.

5.3.1 Theoretical Framework

As mentioned above, the **Appraisal Theory** is best understood as an extension to the broader SFG model. Although a comprehensive account of this model falls outside the immediate scope of this study, a brief overview of its fundamental tenets is necessary. To that end, I will discuss these core principles specifically as they are applied within the context of AD, in relation to its descriptive and evaluative nature.

First and foremost, SFG recognizes language as a context-based carrier of meanings, rather than a rigid classification of prescriptive rules, so that both lexis and grammar are defined by their actual use (Halliday, 1994). Together, they carry out specific functions and realize particular types of meaning (Taylor Torsello, 1996), as determined by context. Context, in turn, guides these functions through the choices that speakers – or, more specifically, describers – make, since they “have reasons for saying something, and for saying it the way they do” (Thompson, 2014: 9). The way we use language is how we construe experience, establish relationships, and produce and share messages. These

three meanings are achieved through the three metafunctions of language – ideational, interpersonal and textual – which are themselves conditioned by contextual factors (Schleppegrell & Oteíza, 2023; Thompson, 2014). Although all the metafunctions operate simultaneously, each is typically accomplished in relation to a specific register variable: the *Field of discourse* (ideational meanings), the *Tenor of discourse* (interpersonal meanings), and the *Mode of discourse* (textual meanings) (Schleppegrell & Oteíza, 2023). In this sense, linguistic choices are a reflection of the context: they define what is happening on screen (*Field*), the role of describers as objective or subjective observers and their relationship with characters (*Tenor*), and the role of language (*Mode*). In particular, in the **ideational** metafunction, the describer uses the lexicogrammatical system of **Transitivity** to encode visual stimuli into linguistic meaning, identifying processes (actions), participants (characters) and circumstances (space/time) appearing on screen. Instead, in the **interpersonal** metafunction, the describer chooses specific **Moods** and **Modalities** systems to present information. For instance, while a declarative mood is used to state facts, modality serves to elicit uncertainty about a character's expression (*epistemic*) or situational pressures upon them (*deontic*). Finally, the **textual** metafunction concerns the relationship between text and images¹⁷. AD, in fact, must be highly cohesive, using **Thematic Organization** (Theme and Rheme) and **Cohesion** to ensure the description fits perfectly within the pauses of the original (cf. Halliday & Matthiessen, 2004; Thompson, 2014; Manfredi, 2011). Taken together, the three metafunctions make up “the set of linguistic resources available to us for making meanings” (Thompson, 2014:39). A practical exemplification is provided below, illustrating how contextual variables (*Field*, *Tenor*, and *Mode*) are realized through lexicogrammatical systems in AD, based on selected samples from *BE*.

¹⁷ SFG also considers a fourth metafunction, closely connected to the ideational: the **logical** metafunction, which explains how events are connected. In AD, logical relations may introduce causal or temporal links that are not explicitly supported by the visual input. This function would require a more detailed analysis, so it will not be discussed in this study.

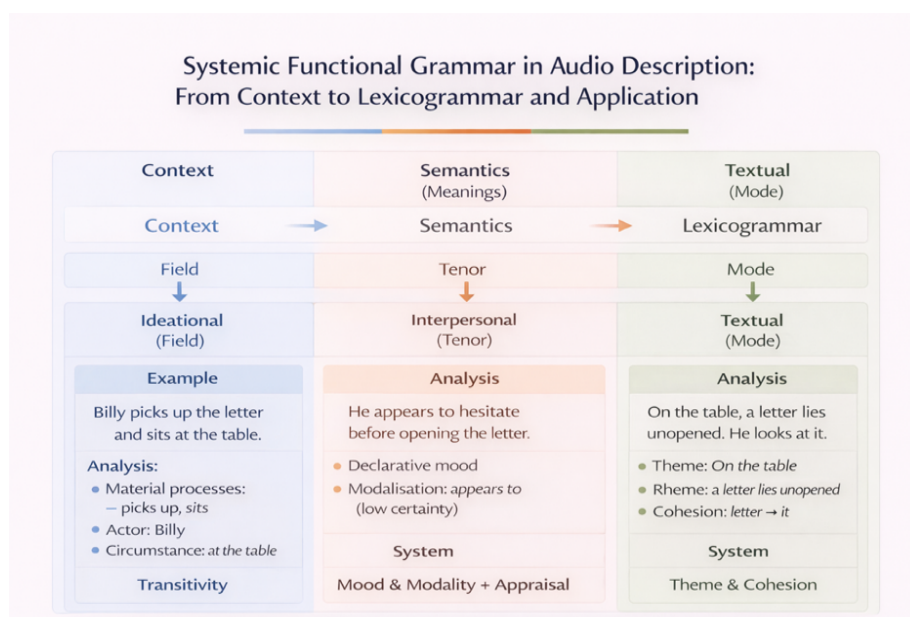


Figure 20. The three metafunctions through AD samples and their lexicogrammatical realizations.

As delineated above, the **Appraisal** system expands upon the interpersonal metafunction as a more recent framework (Martin & White, 2005; cf. Manfredi, 2011) which broadens the traditional account of linguistic resources to incorporate evaluation. Within SFG, Appraisal can be understood as an elaboration of the notion of *Tenor*, capturing not only who is relating to whom, but also how they position themselves towards what is being said, thereby extending interpersonal meaning beyond the basic systems of *Mood* and *Modality*. In the context of AD, this is particularly relevant, as describers do not simply report visual information but may also guide interpretation through evaluative cues.

The Appraisal identifies three systems of meaning that interact with each other: **ATTITUDE**, **ENGAGEMENT** and **GRADUATION**, that are best clarified by referring to an example cited by Rizzo and Pensabene (2021:98), originally sourced from Martin and White (2005:35):

Attitude is concerned with our feelings, including emotional reactions, judgements of behaviour and evaluations of things. Engagement deals with sourcing attitudes and the play of voices around opinions in discourse. Graduation attends to grading phenomena whereby feelings are amplified and categories blurred (Fig. 21).

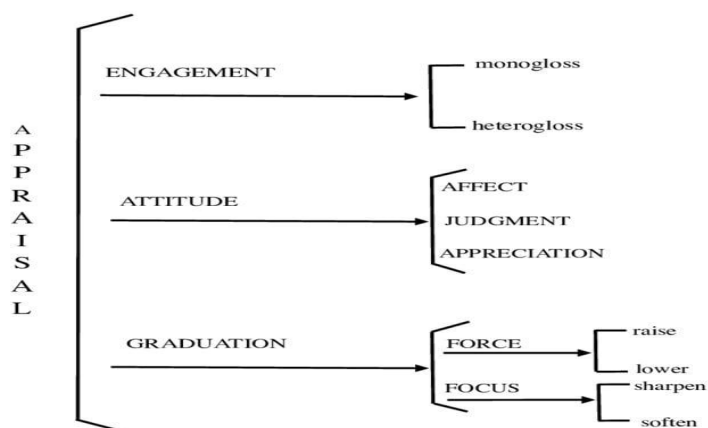


Figure 21: The Appraisal Systems (adapted from Martin & Rose, 2003; cf. Rizzo & Pensabene, 2021).

In the context of AD, when we say that *ATTITUDE* concerns feelings, reactions, and judgements, we are referring to the evaluative significance of AD (*i.e.*, the lexical choices used to describe the emotional state of a character (*AFFECT*), the ethical dimension of an action (*JUDGEMENT*), or the aesthetic value of human and non-human features (*APPRECIATION*). Instead, if we claim that *ENGAGEMENT* addresses the "sourcing" and "play of voices" we are looking at the extent to which a describer presents a visual interpretation either as an absolute, objective fact (e.g., "*he is panicked*") or a subjective impression (e.g., "*he appears panicked*"). Lastly, *GRADUATION* concerns "grading phenomena" by modulating the degree of evaluations, thereby allowing the describer to amplify or mitigate the intensity of an emotion (e.g., "*upset*" vs. "*furious*") as well as adjust the focus of visual categories (e.g., "*emerald green*" vs. "*greenish*") (see Rizzo & Pensabene, 2021).

Central to this discussion is not so much how learners "negotiate evaluations with their addressees" (Thompson, 2014:80) as, rather, how they manage the emotional and evaluative load of the clips. For this reason, I will only take into account the core system of *ATTITUDE* with its related sub-domains: *AFFECT*, *JUDGEMENT* and *APPRECIATION*. To put it simply, Thompson (2014:80) introduces the *ATTITUDE* as "an indication of whether the speaker thinks that something (a person, thing, action, event, situation, idea, etc.) is good or bad". Speaking of AD, this concept implies that the describer establishes an intersubjective bond with such elements that functions along a scale of alignment and agreement. In terms of language, these relations are expressed by wordings of emotional responses (*AFFECT*), ethical responses (*JUDGEMENT*), and aesthetic responses (*APPRECIATION*) (Spinzi, 2016; Fig. 22).

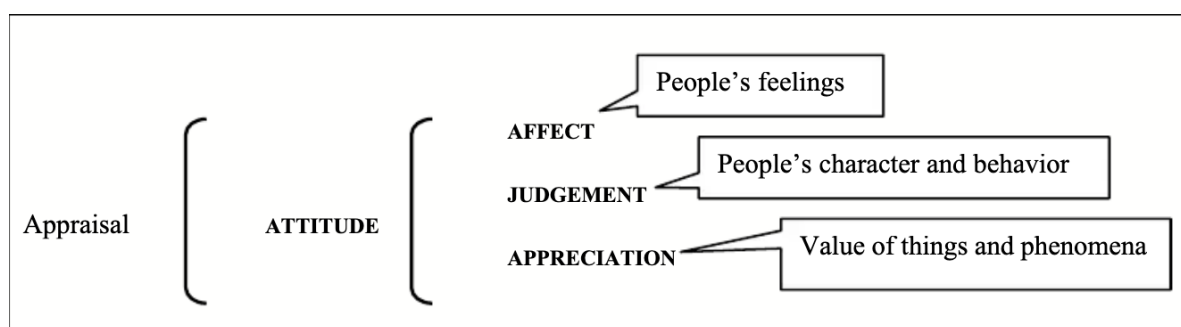


Figure 22: The Attitude system and its three sub-systems based on Martin & Rose (2003) and adapted by Lipson (2006).

As shown above, the system of *AFFECT* provides the grammatical tools to express how we feel about things. These represent a primary method for stance-taking, as they allow the describer to register an emotional response to a phenomenon through mental processes (*He longs for an answer*), attributive relations (*She is relieved*), and minimalization (*Her relief is visible*).

The *JUDGEMENT* sub-system constructs moral evaluations of human behavior, expressing either *social sanction* or *social esteem*; *social sanction* encompasses two categories: *propriety* and *veracity*, both signaling whether an action is legally or ethically culpable (e.g., *He acts fairly*; *he is honest*). *Social esteem* is connected to judgements of *normality* (how customary the behavior is), *tenacity* (how determined the person is), and *capacity* (how capable they are) (Rizzo & Pensabene, 2021; Iedema et al., 1994), e.g., *He behaves oddly*; *she struggles with the task*; *she gives up easily*.

The third sub-system, *APPRECIATION*, refers to how things, phenomena and humans are valued according to social values and aesthetics (Lipson, 2006; Pensabene, 2023). These entities are “appreciated” for their *composition* (e.g., *harmonious*, *symmetrical*) and for the *reaction* with which they are associated, that is their aesthetic impact (e.g., *boring*, *beautiful*, *amazing*). In addition to these, *valuation* concerns the assessment of the social or cultural significance of entities and processes, allowing describers to construe their perceived worth or relevance within a given context (e.g., *crucial*, *noteworthy*, *dangerous*) (Rothery & Stenglin, 2000). More importantly, each subsystem encodes evaluative meaning through positive and negative polarity (+ or –), allowing describers to position their stance towards people, emotions, and phenomena. The following scheme provides an overview of these categories, illustrating how *AFFECT*, *JUDGEMENT*, and *APPRECIATION* are realized through positive and negative values.

Positive (+) and negative (–) evaluative meanings of the ATTITUDE sub-system

Affect (feelings)

- +: joy, love, interest, confidence
 - –: sadness, fear, anger, boredom
-

Judgement (evaluation of behavior)

Social Esteem

– Normality

- + ordinary, typical
- – odd, eccentric

– Capacity

- + skilled, capable
- – incompetent, slow

– Tenacity

- + determined, reliable
- – lazy, unreliable

Social Sanction

– Veracity

- + honest, truthful
- – deceitful, dishonest

– Propriety

- + ethical, fair
 - – immoral, unjust
-

Appreciation (evaluation of things, processes)

– Reaction

- + engaging, appealing
- – boring, dull

– Composition

- + balanced, clear
- – chaotic, unclear

– Valuation

- + important, valuable
- – trivial, worthless

In the context of this study, Appraisal Theory serves as a tool to discern how learners interpret audiovisual elements and, more specifically, how they construct and express varying degrees of subjectivity through their evaluative stance in the descriptive process. All that said, the following section moves to examine the extent to which the ATTITUDE resources appear in the descriptions mostly as *classifiers*, *attitudinal epithets*, *relational attributes* and *processes*, and *prepositional phrases*, as well as *nominal groups*.

5.3.2 Analysis of Audio Description segments through the Attitude System

This analysis builds on the ATTITUDE sub-system outlined above. More specifically, the presence, type, and intensity of AFFECT, JUDGEMENT, and APPRECIATION resources are used to examine the extent to which learners offered an evaluative appraisal of the audiovisual text. In this sense, evaluation is taken as an indicator of subjectivity, reflecting the describer's intervention beyond neutral reporting. The ATTITUDE system and the notion of subjectivity are, in fact, closely related to each other, as the former provides the linguistic resources through which the latter can be expressed and analyzed. In other words, subjectivity refers to the describer's interpretive involvement in the verbal representation of the AVST and Appraisal accounts for the specific lexico-grammatical resources – AFFECT, JUDGEMENT, and APPRECIATION – through which this involvement is manifested.

As mentioned above, subjectivity is recognized through the identification and classification of evaluative resources by type and levels of intensity, but rather than a binary category, it is a scalar phenomenon that exists along a continuum. To this end, the scripts are positioned along an appraisal-based subjectivity continuum, ranging from minimal or absent evaluation to highly marked evaluative stance, expressed by positive or negative polarity ($\pm$) (e.g., (+) engaging, (–) terrified, etc.). At the lowest end, descriptions with no ATTITUDE are classified as highly objective; a first level of subjectivity emerges through implicit or low-level APPRECIATION, where evaluation is still subtle and often inferable. Moderate subjectivity is associated with explicit APPRECIATION, while higher levels of

subjectivity involve the combined use of APPRECIATION with JUDGEMENT and/or AFFECT, hinting at a more interpretive engagement with the AVST. Finally, strong subjectivity is characterized by the presence of marked AFFECT and JUDGEMENT, thus accentuating the describer’s bias. Through this analytical framework, it becomes possible to capture the degrees of ATTITUDE, through which we can determine not only the extent of evaluation, but also how the learners position themselves towards a given situation.

In light of this, the analysis draws on a fresh sample of AD – three per genre – all of which are distinct from those analyzed in Section 5.1. For the most part, its interpretation is based on a “cumulative groove”, rather than single words, with evaluation construed over small stretches of text (Spinzi, 2016; Miller, 2007). The subjectivity continuum below is used for positioning the segments according to the type, polarity and intensity of evaluative resources. Strong subjectivity is distinguished from high subjectivity by the presence of evaluative layering, where emotional interpretation is combined with social and interpretive evaluation (see Table 20 below).

Level	Degree of Subjectivity	Appraisal Resources	Interpretive Profile
1	Objective	No Attitude	Fully denotative; no evaluation or interpretation
2	Low	Implicit Appreciation	Minimal, inferable evaluation
3	Moderate	Appreciation (+/-)	Explicit but limited evaluation of scenes or elements
4	High	Strong Affect (+/-) and/or Appreciation (+/-) + Affect/Judgement	Emotional amplification; interpretation remains affective
5	Strong	Affect (+/-) + Judgement (+/-) (high intensity)	Reconstruction involving emotional, social, and interpretive evaluation

Table 20: Appraisal-based Subjectivity continuum.

Appraisal systems of ATTITUDE

– **Affect (±)**: emotional states

– **Judgement (±)**: evaluation of behavior (*normality; capacity; tenacity; veracity; propriety*)

– **Appreciation (±)**: evaluation of objects, scenes, or processes (*reaction, composition, valuation*)

– Classifier (C); Attitudinal Epithet (AE); Relational Attribute (RA); Prepositional Phrase (PP); Nominal Group (NG); Process (P); Relational Process (RP)

BE

1.

[00:00:45.05] *They're all on the verge of tears, especially his grandmother, who is a nervous wreck.*

Appraisal Resources

AFFECT (–):

- *on the verge of tears* -> high-intensity negative AFFECT (PP)
- *is a nervous wreck* -> high-intensity negative AFFECT (RA)

JUDGEMENT (–) implicit:

- *nervous wreck* -> implies nuanced *capacity*

Continuum Position

High Subjectivity (Strong AFFECT)

Interpretation

The description foregrounds the emotional state of the characters through explicit and highly intensified AFFECT. Resulting from a guided choice introduced as a pedagogical requirement for the participants, expressions such as *on the verge of tears* and *a nervous wreck* – partially inferable from visual cues – index strong emotional responses, bringing psychological distress to the fore.

2.

[00:01:57.05] *Billy finishes the letter and realizes the big news is turning into something that can ruin his life due to his father's disapproval, namely the acceptance at the Royal Ballet School.*

Appraisal Resources

AFFECT (–):

- *ruin his life* -> high-intensity negative AFFECT (P)

JUDGEMENT (–) implicit:

- *can ruin his life* -> implicit *capacity/propriety*
- *father's disapproval* -> implicit *propriety* (NG)

APPRECIATION (+/–):

- *big news* -> low-intensity *valuation* (AE)

Continuum Position

Strong subjectivity (AFFECT + JUDGEMENT)

Interpretation

The description revolves around an interpretive account of the character's situation. The process *can ruin his life* includes high-intensity negative AFFECT, introducing the event in catastrophic terms; instead, the nominal group alludes to social evaluation. Together, these elements are strong subjectivity indicators, through which the student infers emotional states, but also their causes and consequences. This configuration of AFFECT and JUDGEMENT adopts a strongly evaluative stance towards the event.

3.

[00:00:21.07] *The air is very tense and everyone looks at him straight in the eyes, curious to know.*

Appraisal Resources

APPRECIATION (-):

- *very tense* -> medium-high-intensity negative APPRECIATION (RA)

AFFECT (+) implicit:

- *curious to know* -> low-intensity AFFECT (RA)

Continuum Position

High subjectivity (APPRECIATION + AFFECT)

Interpretation

The description blends atmospheric evaluation with an inferred emotional state. The relational attribute *very tense* sets the mood of the scene, while *curious to know* defines the characters' mental state from their behavior. Together, these elements introduce a moderate degree of interpretive intervention. In Appraisal terms, the combination of APPRECIATION and AFFECT, positions the student as a biased interpreter.

ES

4.

[00:00:59.04] *A little girl who has her head stuck in a chair and a pumpkin-shaped box enters the studio.*

Appraisal Resources

APPRECIATION (+/-):

- *pumpkin-shaped box* -> neutral/low-level APPRECIATION (*composition*), very low-intensity (C)

Continuum Position

Low-Moderate subjectivity (low-intensity APPRECIATION)

Interpretation

The description remains largely denotative. The reference to a *pumpkin-shaped box* introduces a minimal degree of evaluative description, but this functions more as a

specification rather than a subjective interpretation. The near absence of ATTITUDE results in a neutral stance.

5.

[00:01:41.11] *The third kid comes in. He seems very shy and jittery.*

Appraisal Resources

AFFECT (–):

- *very shy* -> medium intensity AFFECT (RP)
- *jittery* -> medium intensity AFFECT (RP)

Continuum Position

High subjectivity (AFFECT)

Interpretation

The relational process *seems very shy and jittery* introduces medium-intensity emotional states that, again, are not directly observable but suggested by the characters' behavior. The use of *seems* mitigates the strength of the claim, so that evaluation is more tentative than categorical. However, the presence of a medium degree of ATTITUDE marks a clear tendency to interpret emotions rather than describing them.

6.

[00:00:24.24] *A kid shows up with a grey wig, a white shirt, and beige trousers, holding a pumpkin basket.*

Appraisal Resources

APPRECIATION:

- *grey wig, white shirt, beige trousers (composition)* non-evaluative polarity (AE)
- *pumpkin basket* -> non-evaluative description (C)

Continuum Position

Objective (no ATTITUDE)

Interpretation

The description remains fully denotative, focusing on visual details such as clothing and objects. The lexical choices do not encode elements of ATTITUDE, resulting in a neutral stance.

VITS

7.

[00:01:14.12] *Villagers live in difficult conditions, struggling daily to survive with little means available.*

Appraisal Resources

APPRECIATION (–):

- *in difficult conditions* -> negative APPRECIATION (*valuation*), medium intensity (PP)

AFFECT (–) implicit:

- *struggling daily to survive* -> medium-high intensity AFFECT (PP)

JUDGEMENT (–) implicit:

- *with little means available* -> implicit social evaluation (*capacity/propriety*) (PP)

Continuum Position

Strong subjectivity (AFFECT + JUDGEMENT, reinforced by APPRECIATION)

Interpretation

The expression *difficult conditions* is a clear example of negative APPRECIATION, while *struggling daily to survive* evokes AFFECT as a response to the hardship of drought. The phrase *with little means available* adds social evaluation, suggesting limited resources. Overall, this configuration shows the student's affective proximity with socio-economic hardship.

8.

[00:00:48.13] *Close-up of a battered house in the village; the roads are rocky and everything looks ruined.*

Appraisal Resources

APPRECIATION (–):

- *battered house* -> negative APPRECIATION (*valuation*), medium intensity (AE)
- *rocky roads* -> negative APPRECIATION (*composition*), low-medium intensity (AE)
- *everything looks ruined* -> negative APPRECIATION (*valuation*), high intensity (RP)

Continuum Position

Moderate subjectivity (APPRECIATION –)

Interpretation

The student construes the setting as deteriorated through a cluster of negative APPRECIATION resources. Terms such as *battered*, *rocky*, and *ruined* portray the environment as degraded and emphasize deterioration.

9.

[00:01:24.14] *A group of aged people sits on a porch; they seem tired and hopeless.*

Appraisal Resources

AFFECT (–):

- *seem tired* -> negative AFFECT low-medium intensity (RP)
- *(seem) hopeless* -> negative AFFECT high intensity (RP)

Continuum Position

High subjectivity (AFFECT –)

Interpretation

The student infers the group's emotional state with negative *AFFECT*, as *tired* and *hopeless* suggest a feeling of weariness and despair. The evaluation remains an inferential projection, and is therefore inevitably subjective.

5.3.3 Overall Patterns of Subjectivity in the Scripts

A comparative analysis of the selected AD segments reveals a certain predisposition toward evaluative language, often articulated through interpretive renderings within expressive diegetic contexts. In most films, the performers' expressive baseline is often more heightened than real-world prototypes (Bordwell, 2008) and this is certainly true for *BE*, yet not entirely applicable to *VITS*, due to its non-fictional nature. In both cases, though, the “Architecture of Subjectivity” is effectively construed by a number of recurrent negative attitudinal resources, particularly *AFFECT* and *APPRECIATION*.

However, the intensity of this subjectivity varies by genre. The descriptions for the *ES* segments show moderate levels of subjectivity falling under *APPRECIATION*, although this is not surprising if we think that the aim was to prioritize visual attributes (e.g., appearance and objects) that may reduce interpretive expansion.

The patterns identified in the analysis are summarized in Table 21, which illustrates the distribution of subjectivity levels and *APPRAISAL* resources.

Dataset	Typical Subjectivity Level	Dominant Appraisal Resources	Interpretive Tendencies
Billy Elliot	High → Strong	Affect (–), + occasional Judgement	Emotional inference; focus on characters' internal states
Ellen	Objective → Moderate	Appreciation (+/–)	Visual description; focus on observable details; limited interpretive expansion
Aagaswadi	Moderate → Strong	Appreciation (–) + Affect (–) + Judgement (–)	Socio-economic evaluation; generalisation; interpretive and explanatory description

Table 21. Overview of subjectivity patterns across the three clips.

These patterns indicate that learners adopt different evaluative stances depending on the audiovisual text. In more emotionally or socially complex contexts, they position themselves as interpreters who create meaning through attitudinal resources, whereas in a studio-based context, they tend to adopt a more observational stance, adapting their descriptions to the timing of visual gags and preserving the general rhythm.

In conclusion, we may say that subjectivity is closely linked to creative engagement in AD. The use of ATTITUDE resources reflects the learners' willingness to interpret – where possible – and even expand the visual cues, drawing on their linguistic proficiency to foster creativity.

Concluding Remarks

This study has traced the evolutionary path of filmic AD from a compensatory modality for the visually impaired into an innovative and multifaceted tool. The effects produced by universalist and proactive accounts of accessibility have triggered the creative potential of AD, crossing the boundaries of VI to benefit sighted audiences in various contexts, most notably in educational settings.

Following the recent updates to the CEFR/CV (2020), DAT has been integrated into FLE as a powerful methodology complementing long-established learning programs. Unlike other audiovisual modalities – such as subtitling and dubbing – AD has only recently gained recognition in the language classroom, prompting teachers to design specific activities and adapt them to learners' proficiency levels. Not only has the implementation of task-based exercises shown positive effects on learners' reception, production, mediation, and interaction skills; it has also proven beneficial to their intercultural knowledge, critical and creative thinking. The core objective of this research was to test the pedagogical value of CADAD for the development of written production in L2.

CADAD represents far more than an innovative classroom activity. Rather, it constitutes an interdisciplinary pedagogical framework where AS, TS, Applied Linguistics, AVT, Multimodality, and FLE converge to redefine both the teaching of writing and the role of accessibility within educational contexts. By positioning learners as mediators of multimodal meaning rather than passive recipients of language, CADAD challenges traditional conceptions of writing as a purely linguistic exercise and instead promotes it as an act of interpretation, mediation, creativity and communicative responsibility.

The findings – namely, the learners' lexical enrichment, their resolution of intersemiotic challenges, and the structuring of an evaluative stance – suggest that accessibility should no longer be understood exclusively as a compensatory practice designed for specific audiences. In line with universalist approaches to MA, the research supports the view that accessibility can become a pedagogical principle capable of benefiting all learners. Viewed through this lens, AD becomes a powerful educational resource through which students develop lexical sophistication, multimodal awareness, interpretive competence and audience sensitivity. Accessibility, therefore, ceases to be the object of instruction and becomes the methodology through which FLE itself takes place.

From a theoretical perspective, this research contributes to the growing field of DAT by proposing CADAD as a distinct pedagogical model grounded in the AoA and informed by functionalist TS. The study further demonstrates that creativity in AD should not be conceived as unrestricted subjective expression, but rather as purposeful communicative decision-making guided by the function of the text, the needs of its audience, and the translator's responsibility as an intercultural mediator. In this sense, creativity and accessibility are not opposing forces but complementary dimensions of effective communication.

The corpus analysis further highlights how engaging with multimodal texts through CADAD encourages learners to negotiate meaning at multiple levels simultaneously. Throughout the writing process, participants did not simply translate visual information into words. They interpreted emotions, selected relevant details, evaluated communicative priorities, and continuously balanced descriptive precision with expressive language. Such processes reveal the cognitive, multimodal, and reflective nature of AD, confirming its value as a language-learning activity capable of fostering higher-order thinking alongside linguistic competence.

Beyond language development, the pedagogical implications of this research are equally significant. The CADAD tasks encouraged learners to become active social agents who experienced authentic communicative situations resembling professional accessibility practices. By working within a cloud-based professional environment and engaging with genuine audiovisual materials, learners developed not only writing skills but also an increased awareness of inclusion, accessibility, and the ethical implications of communication. In this respect, CADAD contributes to the development of what may be defined as accessibility literacy: the ability to recognize communicative barriers, reflect critically upon them and make informed linguistic choices that facilitate access for diverse audiences.

Importantly, adopting this framework requires learners to balance pedagogy-oriented expressive freedom with the constraints of professional requirements. This necessity arises because subjective and creative language promotes richer L2 vocabulary acquisition, whereas professional AD prioritizes factual objectivity and non-intrusiveness. Thus, students can be taught to navigate this boundary through a scaffolded pedagogical approach: initial tasks can encourage exploratory, open-ended drafting to foster lexical richness and interpretive complexity, followed by peer-review and revision phases guided by professional

standards (e.g., ADLAB or Ofcom guidelines). In this structured progression, teachers can provide rubrics and analytical checklists to facilitate the transition from creative expression to normative refinement, gradually phasing out this support as students assimilate professional constraints. This dual process allows learners to critically bridge the gap between expressive freedom and professional accuracy, acquiring both linguistic confidence and audience-oriented compliance.

The interdisciplinary nature of this study also deserves emphasis, as briefly outlined at the beginning of the concluding remarks. Accessibility today can no longer be viewed solely through the lens of disability or assistive technologies. Instead, it emerges as a dynamic field that intersects translation, education, linguistics, multimodality, media studies, digital technologies, and cognitive processes. By bringing these disciplines into dialogue, the present research illustrates how innovation in FLE increasingly arises from crossing disciplinary boundaries rather than remaining confined within them. In brief, CADAD exemplifies this convergence, demonstrating that inclusive communication and language learning can mutually reinforce one another.

The thesis, divided into five parts, illustrated the following aspects. In Chapter 1, focus was on establishing the theoretical and historical evolution of MA and AS, highlighting the paradigm shift toward proactive, universalist design and AFM frameworks. In Chapter 2, attention turned to the developmental trajectory of AD as an academic and professional discipline, starting from clinical views of blindness, to the rise of creative applications of the practice, and its pedagogical value within FLE to boost L2 writing skills. Subsequently, Chapter 3 detailed a practical classroom test conducted at the University of Palermo, where B2-level Italian upper-secondary learners of English engaged in a CADAD workshop. Working with three distinct audiovisual genres on the cloud-based *Frazier® - VideoToVoice* platform, the participants were encouraged to prioritize interpretive, vivid language over strict objectivity, while the source texts were categorized from a TS perspective to analyze their intersemiotic challenges and select the most suitable descriptive strategies.

In Chapter 4, the research transitioned into an analytical phase, where a software-aided quantitative investigation of the genre-specific corpus was followed by an analysis of the emerging lexical priorities inspired by the ADLAB guidelines. This chapter further established a task-based assessment framework by adapting the CEFR/CV (2020) descriptors to evaluate intralingual AD as a tool for text mediation and creative writing production. Finally, the fifth Chapter offered a qualitative research perspective, exploring

the subjective value that learners attached to the visual elements they described. The analysis drew from a taxonomy combining Bardini's (2020) ADTs and Rizzo and Spinzi (2023)'s strategies along a "creativity continuum", followed by the *ATTITUDE* sub-system of the Appraisal Theory, determining how learners constructed an evaluative stance and expressed subjectivity through specific language structures.

Naturally, the study presents certain limitations. The experimental design involved a relatively homogeneous group of Italian upper-secondary learners of English and focused on three audiovisual genres within a written production task. Although these choices ensured methodological consistency, further research should investigate the applicability of CADAD across different educational levels, linguistic backgrounds, proficiency ranges and cultural contexts. Longitudinal studies may also explore the long-term impact of repeated exposure to CADAD on learners' lexical development, creative writing abilities, mediation competences and accessibility awareness.

Moreover, building upon Professor Micol Beseghi's research on the use of Gen-AI in DAD (Section 2.7), future developments in multimodal AI and automated image-to-text tools are expanding the horizons of accessible media production. Still, despite their ability to rapidly produce preliminary descriptive outputs, AI-powered systems cannot replicate the cultural sensitivity, stylistic judgment, and pedagogical purpose essential to authentic mediation. Because of such limitations, human-mediated CADAD does not become obsolete; rather, the scaffolded approach developed in this study elevates the scope of learners' authorship by changing their role from mere drafting agents into critical editors who must evaluate, humanize, and refine machine-generated texts against professional standards. Consequently, understanding this balance between human creativity and AI within CADAD promises to open up a highly rewarding direction for further investigation in both MA and FLE.

Ultimately, the significance of this work extends beyond the teaching of English writing. It argues for a broader reconceptualization of accessibility as a pedagogical philosophy capable of transforming the language classroom into a space where multimodal literacy, creativity, inclusion, and intercultural mediation develop simultaneously. In doing so, it positions AD not simply as a technical utility, but as a socially meaningful communicative practice that invites learners to engage critically with images, language, and audiences alike. Far from being a mere tool for content adaptation, AD operates as a transformative mode of mediation that connects languages, modes, disciplines and

communities. Through CADAD, learners do not merely acquire linguistic competence: they learn to observe more attentively, interpret more critically, write more creatively and communicate more responsibly. In an increasingly multimodal and inclusive society, these abilities are not ancillary to language education but are among its most essential objectives.

Bibliography

- AENOR. (2005). *UNE 1530220. Audiodescripción para personas con discapacidad visual. Requisitos para la audiodescripción y elaboración de audioguías.*
- American Academy of Ophthalmology. (2025, November 25). Vitamins for AMD. Retrieved October 14, 2025, from <https://www.aao.org/eye-health/diseases/vitamins-amd>
- American Council of the Blind. (2003). *Guidelines for Audio Describers*. The Audio Description Project. Retrieved August 15, 2023, from <https://adp.acb.org/guidelines.html>
- Androutsopoulos, J. (2012). Introduction: Language and society in cinematic discourse. *Multilingua*, 31(2-3), 139-154. <https://doi.org/10.1515/multi-2012-0007>
- Anthony, L. (2004). AntConc: A learner and classroom friendly, multi-platform corpus analysis toolkit. *Proceedings of IWLeL 2004: An Interactive Workshop on Language e-Learning*. IWLeL 2004 Program Committee. Retrieved April 13, 2025, from https://www.researchgate.net/publication/33010037_AntConc_A_Learner_and_Classroom_Friendly_Multi-Platform_Corpus_Analysis_Toolkit
- Anthony, L. (2005). *AntConc: design and development of a freeware corpus analysis toolkit for the technical writing classroom* [Conference presentation]. IPCC 2005. Proceedings. International Professional Communication Conference, Limerick.
- Anthony, L. (2006). Developing a Freeware, Multiplatform Corpus Analysis Toolkit for the Technical Writing Classroom. *IEEE Transactions on Professional Communication*, 49(3), 275-286.
- Anthony, L. (2013). A critical look at software tools in corpus linguistics. *Linguistic Research*, 30(2), 141-161.

Austin, J. L. (1970). Performative utterances. In J. L. Austin, *Philosophical papers* [Original work transcribed from a BBC Third Programme broadcast, 1956] (pp. 233-252). Oxford University Press.

Baker, M. (1992). *In other words: A coursebook on translation*. Routledge.

Baker, M. (2006). Translation and context. *Journal of Pragmatics*, 38(3), 317-320.

Baker, M., & Saldanha, G. (2009). *Routledge encyclopedia of translation studies* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203872062>

Baranova, J., & Duoblienė, L. (2019). Multimodal strategies in teaching ethics with films. *Acta Paedagogica Vilnensia*, 43, 71-84. <https://doi.org/10.15388/ActPaed.43.5>

Barbot, B., Tan, M., Randi, J., Santa-Donato, G., & Grigorenko, E. L. (2012). Essential skills for creative writing: Integrating multiple domain-specific perspectives. *Thinking Skills and Creativity*, 7(3), 209-223. <https://doi.org/10.1016/j.tsc.2012.04.006>

Bardini, F. (2020a). Audio description and the translation of film language into words. *Ilha do Desterro – A Journal of English Language Literatures in English and Cultural Studies*, 73(1), 273-296. <https://doi.org/10.5007/2175-8026.2020v73n1p273>

Bardini, F. (2020b). Film language, film emotions and the experience of blind and partially sighted viewers: a reception study. *The Journal of Specialised Translation*, 33, 259-280. <https://doi.org/10.26034/cm.jostrans.2020.556>

Bardini, F. (2022). Audio description, subjectivity and film experience. *Quaderns del CAC*, 48, 97-106.

Bausells-Espín, A. (2022). Audio description as a pedagogical tool in the foreign language classroom: An analysis of student perceptions of difficulty, usefulness and learning progress. *Journal of Audiovisual Translation*, 5(2), 152-175. <https://doi.org/10.47476/jat.v5i2.2022.208>

Bausells-Espín, A. (2024). Student perceptions of transferable skills development through didactic audio description. *Cadernos de Tradução*, 44(1). <https://doi.org/10.5007/2175-7968.2024.e94332>

Beseghi, M. (2026, April 23-24). *From words to images: Using AI-generated visuals in didactic audio description* [Conference presentation]. Artificial Intelligence and Audiovisual Translation, University of Palermo.

Bestard-Bou, J., & Arias Badia, B. (2022). Access services for the blind and partially sighted. A social and legal framework for the promotion of audio description. In C. Taylor & E. Perego (Eds.), *The Routledge handbook of Audio Description* (pp. 27-37). London and New York: Routledge.

Betz, M. (2001). The name above the (sub)title: Internationalism, coproduction, and polyglot European art cinema. *Camera Obscura*, 46, 1-44.

Bittner, H. (2012). Audio description guidelines: A comparison. *New Perspectives in Translation*, 20, 41-61.

Bondebjerg, I. (2015). Film: Genres and genre theory. *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed.), 160-164. <https://doi.org/10.1016/B978-0-08-097086-8.95052-9>

Bordwell, D. (2008). *Poetics of cinema* (1st ed.). Routledge. <https://doi.org/10.4324/9780203941898>

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 J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for All: Subtitling for the Deaf, Audio Description, and Sign Language* (pp. 175-187). Rodopi.

Branson, J. (2024). *Accessible filmmaking in practice: critical and ethnographic perspectives on inclusionism, errancy and excess* [Doctoral dissertation], University of

Roehampton.

[https://pure.roehampton.ac.uk/ws/portalfiles/portal/26403858/joshua branson accessible filmmaking in practice thesis.pdf](https://pure.roehampton.ac.uk/ws/portalfiles/portal/26403858/joshua_branson_accessible_filmmaking_in_practice_thesis.pdf)

Braun, S. (2008). Audio Description research: State of the art and beyond. *Translation Studies in the New Millennium*, 6, 14-30.

Braun, S., & Starr, K. (2021). *Innovation in Audio Description Research*. Routledge.

Brechin, A., Liddiard, P., & Swain, J. (1981). *Handicap in a social world: A reader*. Hodder and Stoughton.

Brescia-Zapata, M., & McDonagh, S. (2024). Exploring the potential of media accessibility as a pedagogical tool: Students as audio describers. In A. Marcus-Quinn, K. Krejtz, & C. Duarte (Eds.), *Transforming Media Accessibility in Europe* (pp. 317-336). Springer, Cham. https://doi.org/10.1007/978-3-031-60049-4_18

Broderick, A. (2020). Of rights and obligations: the birth of accessibility. *The International Journal of Human Rights*, 24(4), 393-413. <https://doi.org/10.1080/13642987.2019.1634556>

Brown, W. (2016b). Non-cinema: Digital, ethics, multitude. *Film-Philosophy*, 20(1), 104-130. <https://www.euppublishing.com/doi/10.3366/film.2016.0006>

Buckingham, D. (n.d.). *Ratcatcher*. Growing Up Modern. Retrieved November 24, 2025, from <https://davidbuckingham.net/growing-up-modern/what-can-a-poor-boy-do/ratcatcher/>

Cabrera, L. (2009). Nanotechnology: Changing the disability paradigm. *International Journal of Disability, Community & Rehabilitation*, 8(2).

Calduch, C., & Talaván, N. (2018). Traducción audiovisual y aprendizaje del español como L2: el uso de la audiodescripción. *Journal of Spanish Language Teaching*, 4(2), 168-180.

Cámara, L., & Espasa, E. (2011). The audio description of scientific multimedia. *The Translator*, 17(2), 415-437. <https://doi.org/10.1080/13556509.2011.10799496>

Carradini, S., & Swarts, J. (2023). *Text at scale: Corpus analysis in technical communication*. The WAC Clearinghouse. University Press of Colorado. <https://doi.org/10.37514/TPC-B.2023.2104>

Cavallo, A. (2015). Seeing the word, hearing the image: the artistic possibilities of audio description in theatrical performance. *Research in Drama Education: The Journal of Applied Theatre and Performance*, 20(1), 125-134. <http://dx.doi.org/10.1080/13569783.2014.983892>

Chaume, F. (2004). *Cine y Traducción* [Cinema and Translation]. Cátedra.

Chaume, F. (2013). The turn of audiovisual translation. New audiences and new technologies. *Translation Spaces*, 2(1), 105-123.

Chaume, F. (2018). Is audiovisual translation putting the concept of translation against the ropes? *The Journal of Specialised Translation*, 30(1), 85-104.

Chmiel, A., & Mazur, I. (2011). Audiodeskrypcja jako intersemiotyczne przekład audiowizualny – percepcja produktu i ocena jakości. In I. Kasperska & A. Żuchelkowska (Eds.), *Przekład jako produkt i kontekst jego odbioru* (pp. 13-30). Wydawnictwo Rys.

Chmiel, A., & Mazur, I. (2012). AD reception research: Some methodological considerations. In E. Perego (Ed.), *Emerging topics in translation: Audiodescription* (pp. 57-80). Trieste: EUT.

Chmiel, A., & Mazur, I. (2016). Researching preferences of audio description users – Limitations and solutions. *Across Languages and Cultures*, 17(2), 271-288. <https://doi.org/10.1556/084.2016.17.2.7>

Clarkson, P. J., & Coleman, R. (2015). History of inclusive design in the UK. *Applied Ergonomics*, 46, 235-247. <https://doi.org/10.1016/j.apergo.2013.03.002>

Colmenero, M. O. L. (2024). Subjectivity and creativity versus audio description guidelines. In A. Marcus-Quinn, K. Krejtz, & C. Duarte (Eds.), *Transforming Media Accessibility in Europe* (pp. 39-52). Springer, Cham. https://doi.org/10.1007/978-3-031-60049-4_3

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Council of Europe.

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume with New Descriptors*. Council of Europe.

Cutting, J. E. (2019b). Simplicity, complexity, and narration in popular movies. In M. Grišakova & M. Poulaki (Eds.), *Narrative complexity: Cognition, embodiment, evolution* (pp. 200-222). University of Nebraska Press.

Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education*, 10, 241-266. <https://doi.org/10.1177/1028315306287002>

Díaz Cintas, J. (2005). Audiovisual translation today. A question of accessibility for all. *Translating Today*, 4, 3-5.

Díaz Cintas, J., & Remael, A. (2007). *Audiovisual translation: Subtitling*. Routledge.

Díaz Cintas, J. (2019). Audiovisual translation. In E. Angelone, M. Ehrensberger-Dow, & G. Massey (Eds.), *The Bloomsbury Companion to Language Industry Studies* (pp. 209-230). Bloomsbury.

Díaz Cintas, J., & Remael, A. (2020). *Subtitling: Concepts and Practices*. Routledge.

Di Giovanni, E. (2014). Visual and narrative priorities of the blind and non-blind: eye tracking and audio description. *Perspectives*, 22(1), 136-153.
<https://doi.org/10.1080/0907676X.2013.769610>

Di Giovanni, E. (2018a). Audio description for live performances and audience participation. *The Journal of Specialised Translation*, 29, 189-211.
<https://doi.org/10.26034/cm.jostrans.2018.218>

Di Giovanni, E. (2018b). Participatory accessibility: Creating audio description with blind and non-blind children. *Journal of Audiovisual Translation*, 1(1), 155-169.
<https://doi.org/10.47476/jat.v1i1.50>

Di Giovanni, E. (2020). *La traduzione audiovisiva e i suoi pubblici. Studi di ricezione*. Paolo Loffredo Editore. <https://www.loffredoeditore.com/images/book/collane/la-traduzione-e-i-suoi-pubblici.pdf>

Di Giovanni, E. (2024). Audiovisual translation and media accessibility: new languages, new forms of transfer. *Status Quaestionis*, 27, 357-373.

Di Giovanni, E., Lucchetti, M., Turrini, A., & Raffi, F. (2022). Exploring how accessible cultural practices impact on community development: The case of InclusiveOpera at the Macerata Opera Festival in Italy. *Local Development & Society*, 3, 74-87.
<https://doi.org/10.1080/26883597.2022.2045085>

Dore, M. (2019). Introduction: Exploring the many ways of audiovisual translation. Retranslated, simultaneous, indirect, mediated or what? *Status Quaestionis*, 15, 11-11.
<https://doi.org/10.13133/2239-1983/14568>

Dosch, E., & Benecke, B. (2004). *Wenn aus Bildern Worte werden. Ein Handbuch für Filmbeschreiber*. Bayrischer Blindenbund E. V.

Duff, A. (1989). *Translation*. Oxford University Press.

Dunlop, C., Thurston, M., Firth, K., & Southwood, E. (2025). Experiences of loneliness and isolation among people with visual impairment: A review of literature (2016-2023). *British Journal of Visual Impairment*, 43(3), 719-736.

Edison, T. [The Tommy Edison Experience]. (2012, July 12). *Funny side of being blind: The Tommy Edison Experience trailer/promo* [Video]. YouTube. <https://www.youtube.com/watch?v=eCPjR-94p3U>

Elsaesser, T., & Hagener, M. (2010). *Film theory: An introduction through the senses*. Routledge.

Emmott, C. (1997). *Narrative comprehension. A discourse perspective*. Oxford University Press.

Faula, A., Anggraeni, D., & Wulansari, D. (2022). An analysis basic strategy of verbal humor found in Ellen DeGeneres monologue season 17, 2019: Pragmatic approach study. *Mediova*, 2(1), 59-70.

Finkelstein, V. (1981). To deny or not to deny disability. In A. Brechin, P. Liddiard, & J. Swain (Eds.), *Handicap in a Social World*. OUP/Hodder and Stoughton.

Foerster, A. (2010). Towards a creative approach in subtitling: a case study. In J. Díaz Cintas, A. Matamala, & J. Neves (Eds.), *New insights in audiovisual translation and media accessibility* (pp. 81-98). Rodopi.

Frazier, G. (1975). *The autobiography of Miss Jane Pittman: An all-audio adaptation of the teleplay for the blind and visually handicapped*. San Francisco State University.

Fresno, N. (2017). Approaching engagement in audio description. *RITT (Rivista Internazionale di Tecnica della Traduzione)*, 19, 13-32.

Fryer, L. (2010). Directing in reverse. In I. Kemble (Ed.), *The changing face of translation. Proceedings of the 9th annual Portsmouth translation conference* (pp. 63-72). University of Portsmouth.

Fryer, L. (2016). *An introduction to audio description: A practical guide*. Routledge.

Fryer, L. (2018a). Staging the audio describer: An exploration of integrated audio description. *Disability Studies Quarterly*, 38(3). <https://doi.org/10.18061/dsq.v38i3.6490>

Fryer, L. (2018b). The independent audio describer is dead: Long live audio description! *Journal of Audiovisual Translation*, 1(1), 170-186.

Fryer, L. (2019). Quality assessment in audio description: Lessons learned from interpreting. In E. Huertas-Barros, S. Vandepitte, & E. Iglesias-Fernández (Eds.), *Quality Assurance and Assessment Practices in Translation and Interpreting* (pp. 155-177). IGI Global Scientific Publishing.

Fryer, L., & Freeman, J. (2012a). Presence in those with and without sight: Audio description and its potential for virtual reality applications. *Journal of CyberTherapy & Rehabilitation*, 5(1), 15-23.

Fryer, L., & Freeman, J. (2012b). Cinematic language and the description of film: Keeping AD users in the frame. *Perspectives: Studies in Translatology*, 21(3), 412-426. <https://doi.org/10.1080/0907676X.2012.693108>

Fryer, L. (2021). Accessible filmmaking – a model to follow? In L. Fryer & A. Cavallo (Eds.), *Integrated access in Live Performance*. Routledge.

Fryer, L., & Walczak, A. (2021). Immersion, presence and engagement in audio described material. In S. Braun & K. Starr (Eds.), *Innovation in Audio Description Research* (pp. 13-32). Routledge.

Gambier, Y. (2003). Screen transadaptation: perception and reception. *The Translator*, 9(2), 171-189.

Gambier, Y. (2004). La traduction audiovisuelle: un genre en expansion [Audiovisual translation: a growing field]. *Meta*, 49(1), 1-11. <https://doi.org/10.7202/009015ar>

Geerinck, B., & Vercauteren, G. (2020). Audio describing the mental dimension of narrative characters. Insights from a Flemish case study. *Punctum*, 6, 85-107.

Geering, N. (2019, April 5). *BBC Breakfast News – Rationale method of audio description* [Video]. YouTube. https://www.youtube.com/watch?v=IlzMI5V_wcQ&t=5s

Giansante, L. (2015). Writing verbal description audio tours. *Art Beyond Sight: Museum Education Institute*. www.artbeyondsight.org/mei/verbal-description-training/writing-verbal-description-for-audio-guides/

Gibbs, J. (2002). *Mise-en-scène: film style and interpretation*. Wallflower Press.

Gossett, A., Mansha, M., Barnds, A., & Feidt, D. (2009). Beyond access: A case study on the intersection between accessibility, sustainability, and universal design. *Disability and Rehabilitation: Assistive Technology*, 4(6), 439-450.

Greco, G. M. (2016a). On accessibility as a human right, with an application to media accessibility. In A. Matamala & P. Orero (Eds.), *Researching Audio Description: New Approaches* (pp. 11-34). Palgrave Macmillan.

Greco, G. M. (2016b). On Accessibility as a Human Right, with an Application to Media Accessibility. In A. Matamala, & P. Orero (Eds.), *Researching Audio Description. New Approaches* (pp. 11-34). Palgrave Macmillan.

Greco, G. M. (2018). The nature of accessibility studies. *Journal of Audiovisual Translation*, 1(1), 205-232. <https://doi.org/10.47476/jat.v1i1.51>

Greco, G. M. (2019a). The rise of accessibility studies: Abuses, misuses and the method of poietic design. In C. Stephanidis (Ed.), *HCI International 2019 - Late Breaking Papers* (pp. 15-27). Springer.

Greco, G. M. (2019b). Towards a pedagogy of accessibility: The need for critical learning spaces in media accessibility education and training. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 23-46. <https://doi.org/10.52034/lanstts.v18i0.518>

Greco, G. M. (2022). The question of accessibility. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 13-26). Routledge.

Greco, G. M., & Jankowska, A. (2020). Media accessibility within and beyond audiovisual translation. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 57-81). Palgrave Macmillan.

Greening, J., & Rolph, D. (2007). Accessibility: Raising awareness of audio description in the UK. In J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for all: Subtitling for the deaf, audio description, and sign language* (pp. 127-138). Rodopi.

Griffith, J. (2020, January 30). Ellen DeGeneres surprises black teen told to cut dreadlocks with \$20,000 scholarship. *NBC News*. <https://www.nbcnews.com/news/us-news/ellen-degeneres-surprises-black-teen-told-cut-dreadlocks-20-000-n1125801>

Haasch, P., & Cohen, R. (2022, September 23). Viral child star Greyson Chance says 'manipulative' Ellen DeGeneres 'completely abandoned' him after he appeared on her show. *Business Insider*. <https://www.businessinsider.com/greyson-chance-says-ellen-degeneres-completely-abandoned-viral-child-star-2022-9>

Habicht, J., Viswanathan, S., Carrington, B., Huser, T. U., Harper, R., & Rollwage, M. (2024). Closing the accessibility gap to mental health treatment with a personalized self-referral chatbot. *Nature Medicine*, 30(2), 595-602. <https://doi.org/10.1038/s41591-023-02766-x>

Halliday, M. A. K. (1994). *An Introduction to Functional Grammar*. Edward Arnold.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An Introduction to Functional Grammar* (3rd ed.). Arnold.

Haltom, T. M., & Worthen, M. G. (2014). Male ballet dancers and their performances of heteromascularity. *Journal of College Student Development*, 55(8), 757-778.

Harmon, L. (2019). Translation strategies, techniques, and equivalences in critical approach. *Explorations: A Journal of Language and Literature*, 7, 7-16.

Haskell, I. (2018, April 16). *Rhetorical Analysis: Ellen DeGeneres and Comedy*. Medium. <https://medium.com/@isaac.haskell/rhetorical-analysis-ellen-degeneres-and-comedy-a614e254586>

Hayhoe, S. (2018, August 21). *Flipping descriptions: a new phase of democratising audio description*. Vocaleyes. <https://vocaleyes.co.uk/flipping-descriptions-a-new-phase-of-democratising-audio-description/>

Heffernan, J. A. W. (2015). Notes Toward a Theory of Cinematic Ekphrasis. In S. Ercolino, M. Fusillo, M. Lino, & L. Zenobi (Eds.), *Imaginary Films in Literature* (pp. 1-17). Brill.

Hermosa-Ramírez, I. (2022). Profiling audio description service providers. A questionnaire-based snapshot. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 295-311). Routledge.

Hernández-Bartolomé, A., & Mendiluce-Cabrera, G. (2009). How can images be translated? Audio description, a challenging audiovisual and social gap-filler. *Hermēneus: Revista de Traducción e Interpretación*, 11, 161-186.

Herrero, C., & Escobar, M. (2018). A pedagogical model for integrating film education and audio description in foreign language acquisition. *Translation and Translanguaging in Multilingual Contexts*, 4(1), 30-54. <https://doi.org/10.1075/ttmc.00003.her>

Holsanova, J. (2016). A cognitive approach to audio description. In A. Matamala & P. Orero (Eds.), *Researching audio description* (pp. 49-73). Palgrave Macmillan.

Holsanova, J. (2021). Audio description of art: The role of mental imagery and embodiment. *Lund University Cognitive Studies*, 181, 1-19.

Holsanova, J. (2022). A cognitive approach to audio description – Production and reception processes. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 57-77). London and New York: Routledge.

Holst, J. L. F. (2010). *Creativity in translation – A study of various source and target texts* [Master's thesis] Aarhus School of Business.

Hughes, C., Orero, P., & Rai, S. (2021). Towards a user specification for immersive audio description. In S. Braun & K. Starr (Eds.), *Innovation in Audio Description Research* (pp. 121-134). Routledge.

Hutchinson, R. (2020, October 21). *Audio Description in an Inclusive Museum*. Vocaleyes. <https://vocaleyes.co.uk/audio-description-in-an-inclusive-museum/>

Hyks, V. (2005). Audio description and translation: Two related but different skills. *Translating Today*, 4, 6-8.

Ibáñez Moreno, A., & Vermeulen, A. (2013). Audio description as a tool to improve lexical and phraseological competence in foreign language learning. In D. Tsagari & G. Floros (Eds.), *Translation in Language Teaching and Assessment* (pp. 41-65). Cambridge Scholars Publishing.

Ibáñez Moreno, A., & Vermeulen, A. (2015). Profiling a MALL app for English oral practice. A case study. *Journal of Universal Computer Science*, 21(10), 1339-1361.

Ibáñez Moreno, A., & Vermeulen, A. (2016). VISP: A MALL-based app using audio description techniques to improve B1 EFL students' oral competence. In E. Martin-Monje,

I. Elorza, & B. García-Riaza (Eds.), *Technology-enhanced language learning for specialized domains: Practical applications and mobility* (pp. 266-276). Routledge.

Ibáñez Moreno, A., & Vermeulen, A. (2017). The ARDELE Project: Controlled empirical research on audio description as a didactic tool to improve (meta)linguistic competence in foreign language teaching and learning. In J. Díaz Cintas, & K. Nikolic (Eds.), *Fast-forwarding with audiovisual translation* (pp. 195-211). Multilingual Matters.

Iedema, R., Feez, R., & White, P. R. R. (1994). *Media Literacy*. Disadvantaged Schools Program, NSW Department of School Education.

Igareda, P. (2011). The audio description of emotions and gestures in Spanish-spoken films. In A. Serban, A. Matamala, & J.-M. Lavour (Eds.), *Audiovisual translation in close-up: practical and theoretical approaches* (pp. 223-238). Peter Lang.

Igareda, P., & Matamala, A. (2012). Variations on the Pear Tree experiment: different variables, new results? *Perspectives*, 20(1), 103-123.
<https://doi.org/10.1080/0907676X.2011.632684>

Incalcaterra McLoughlin, L., Lertola, J., & Talaván, N. (Eds.). (2018). *Translation and translanguaging in multilingual contexts. Special issue audiovisual translation in applied linguistics: Educational perspectives*. John Benjamins.

Independent Television Commission. (2000). *ITC Guidance on Standards for Audio Description*.

Jakobson, R. (2012). On linguistic aspects of translation. In L. Venuti (Ed.), *The Translation Studies Reader* (pp. 126-131). Routledge. (Original work published 1959)

Jankowska, A. (2021). Audio describing films: A first look into the description process. *The Journal of Specialised Translation*, 36, 26-52.

Jankowska, A. (2022). Audio description and culture-specific elements. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 107-126). London and New York: Routledge.

Jernigan, K. (1993a). *Resolution 93-01*. The National Federation of the Blind Annual Convention, Dallas, Texas [Conference presentation]. <https://disability-studies.leeds.ac.uk/wp-content/uploads/sites/40/library/bolt-JVIB1-Archive.pdf>

Jernigan, K. (1993b). *The pitfalls of political correctness*. National Federation of the Blind. <https://nfb.org/images/nfb/publications/vod/vodsum0504.htm>

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness*. Cambridge University Press.

Kapsaskis, D. (2024). Creative subtitling as film-transformative practice: From immersion to amazement in Edgar Pêra's *The Baron*. *Meta*, 69(1), 155-173. <https://doi.org/10.7202/1113945ar>

Katan, D. (2008). Translation as intercultural communication. In J. Munday (Ed.), *The Routledge companion to translation studies* (pp. 74-92). Routledge.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology*, 13, 1-12.

Kaufman, J. C., & Glăveanu, V. P. (2021). An overview of creativity theories. In J. C. Kaufman & R. Sternberg (Eds.), *Creativity: An Introduction* (pp. 17-30). Cambridge University Press.

Kecskes, I. (2023). *The Interplay of Linguistic, Conceptual and Encyclopedic Knowledge in Meaning Production and Comprehension. The Socio-Cognitive Approach to Communication and Pragmatics*. Springer, Cham. https://doi.org/10.1007/978-3-031-30160-5_14

Kleege, G. (2016). Audio description described: Current standards, future innovations, larger implications. *Representations*, 135(1), 89-101. <https://doi.org/10.1525/rep.2016.135.1.89>

Kleege, G. (2018). *More than Meets the Eye: What blindness brings to art*. Oxford University Press.

Kleege, G., & Wallin, S. (2015). Audio description as a pedagogical tool. *Disability Studies Quarterly*, 35(2). <https://doi.org/10.18061/dsq.v35i2.4622>

Klishevich, A. (2023). Representing the filmscript: Cinematic ekphrasis in Jonathan Coe's *Mr Wilder and Me* and Billy Wilder's *Fedora*. *Ekphrasis. Images, Cinema, Theory, Media*, 30(4), 55-72. <https://doi.org/10.24193/ekphrasis.30.4>

Kruger, J.-L. (2012). Making meaning in AVT: eye tracking and viewer construction of narrative. *Perspectives*, 20(1), 67-86. <https://doi.org/10.1080/0907676X.2011.632688>

Kurch, A. (2024). Production processes in subtitling for the deaf and hard-of-hearing and audio description: Potentials of partly automated process acceleration with the help of (language) technology. In C. Maaß & I. Rink (Eds.), *Handbook of Accessible Communication*, 15. Frank & Timme.

Kurzon, D. (2007). Towards a typology of silence. *Journal of Pragmatics*, 39(10), 1673-1688. <https://doi.org/10.1016/j.pragma.2007.07.003>

LaPointe, E. A. (2023). Tokenism. In *Ebsco Website*. Retrieved December 9, 2025, from <https://dictionary.cambridge.org/it/dizionario/inglese/taxi>

Legge 3 aprile 2001, n. 138. Classificazione e quantificazione delle minorazioni visive e norme in materia di accertamenti oculistici [Classification and quantification of visual impairments and regulations regarding ophthalmic examinations]. *Gazzetta Ufficiale*, (93). <https://www.parlamento.it/parlam/leggi/01138l.htm>

Lertola, J. (2019). *Audiovisual Translation in the Foreign Language Classroom: Applications in the Teaching of English and Other Foreign Languages*. Research-publishing.net. <https://doi.org/10.14705/rpnet.2019.27.9782490057252>

Lertola, J., & Mariotti, C. (2017). Reverse dubbing and subtitling: Raising pragmatic awareness in Italian ESL learners. *The Journal of Specialised Translation*, 28, 103-121.

Li, M. (2021). *Researching and teaching second language writing in the digital age*. Palgrave Macmillan.

Lipson, M. (2006). *Exploring Functional Grammar*. Clueb.

Manfredi, M. (2011). Systemic Functional Linguistics as a tool for translation teaching: Towards a meaningful practice. *Rivista Internazionale di Tecnica della Traduzione*, 13, 49-62. <https://www.openstarts.units.it/server/api/core/bitstreams/4f71c356-0453-4159-b404-83802dce95ff/content>

Marano, F. (2022). The contemplative style in the ethnographic documentary between observation and immersion. *Visual Ethnography*, 11(2), 39-48. <https://doi.org/10.12835/ve2022.2-112>

Marber, D., & Wilson, A. T. (2024). Imagination as an intellectual virtue. *Analysis*, 85(1), 88-99. <https://doi.org/10.1093/analys/anae051>

Marra, L. (2022). Audio description and multiple levels of meaning in films. A functionalist perspective. *Lingue e Linguaggi*, 53, 303-319.

Marra, L. (2023). The expression of emotions in the Spanish and Italian filmic audio descriptions of The King's Speech. *Journal of Audiovisual Translation*, 6(2), 33-54. <https://doi.org/10.47476/jat.v6i2.2023.247>

Martin, J. R., & White, P. R. R. (2005). *The Language of Evaluation: Appraisal in English*. Palgrave Macmillan.

Massidda, S. (2022). [Review of the book *Subtitling: Concepts and practices*, by J. Díaz Cintas & A. Remael]. *Hikma*, 21(1), 367-370.

Matamala, A., & Orero, P. (2015). Text on screen. In A. Remael, N. Reviers, & G. Vercauteren (Eds.), *Pictures painted into words. ADLAB Audio description guidelines* (pp. 39-42). EUT.

Matamala, A., & Rami, N. (2009). Análisis comparativo de la audiodescripción española y alemana de 'Goodbye Lenin'. *Hermēneus: Revista de Traducción e Interpretación*, 11, 249-266.

Matthiessen, C. M. I. M. (2014a). Applicable discourse analysis. In F. Yan & J. J. Webster (Eds.), *Developing systemic functional linguistics: Theory and application* (pp. 135-205). Equinox.

Matthiessen, C. M. I. M. (2014b). Choice in translation: Metafunctional considerations. In K. Kunz, E. Teich, S. Hansen-Schirra, S. Neumann, & P. Daut (Eds.), *Caught in the middle – language use and translation: A festschrift for Erich Steiner on the occasion of his 60th birthday* (pp. 271-333). Saarbrücken: Saarland University Press.

Mazur, I. (2014). Projekt ADLAB i funkcjonalizm w przekładzie – w stronę strategii audiodeskrypcyjnych [The ADLAB project and functionalism in translation – towards audio description strategies]. *Przekładaniec: A Journal of Translation Studies*, 28, 11-22.

Mazur, I. (2015). Narratological building blocks. In A. Remael, N. Reviers, & G. Vercauteren (Eds.), *Pictures painted into words: ADLAB Audio description guidelines* (pp. 19-23). Trieste: EUT.

Mazur, I. (2020a). A functional approach to audio description. *Journal of Audiovisual Translation*, 3(1), 226-245.

Mazur, I. (2020b). Audio description: Concepts, theories and research approaches. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media*

Accessibility (pp. 227-248). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-42105-2_12

Mazur, I. (2022). Linguistic and textual aspects of audio description. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 93-106). Routledge.

Mazur, I. (2024). Same film, different audio descriptions: describing foreign films from a functional perspective. *Universal Access in the Information Society*, 23, 511-521. <https://doi.org/10.1007/s10209-023-00998>

Mazur, I., & Chmiel, A. (2021). Audio description training: A snapshot of the current practices. *The Interpreter and Translator Trainer*, 15(1), 51-65.

Mazzucchelli, L. (2013, August 1). *Il cerchio dei sensi: la vista è davvero così importante?* Psicologo-Milano. <https://www.psicologo-milano.it/newblog/psicologia-cecita-vista/>

McCall, S., McLinden, M., & Douglas, G. (2014). *A review of the literature into effective practice in teaching literacy through braille* [online]. University of Birmingham. https://www.pathstoliteracy.org/wp-content/uploads/2022/06/braille_lit_review.pdf

McQuail, D. (1997). *Audience Analysis*. Sage Publications.

Mennesson, C. (2009). Being a man in dance: socialization modes and gender identities. *Sport in Society*, 12(2), 174-195.

Michalko, R., & Titchkosky, T. (2024). Blindness and dyslexia in the movements of everyday life in Toronto. In S. B. Mintz & G. Fraser (Eds.), *Placing Disability: Literary Disability Studies*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-031-41219-6_9

Miller, D. R. (2007). Towards a typology of evaluation in parliamentary debate: from theory to practice – and back again. In M. Dossena & A. Jucker (Eds.), *(R)evolution in Evaluation*, *Textus*, 20(1), 159-180.

Miller, L. (2023, February 14). *Multilingual technology for global entertainment: A case study of Netflix*. Arts Management & Technology Laboratory. <https://amt-lab.org/blog/2023/1/the-future-of-non-english-storytelling-through-linguistic-diversity-a-case-study-of-netflix>

Ministero dell'Istruzione e del Merito (2019). *Linee guida per i percorsi per le competenze trasversali e per l'orientamento* (Decreto Ministeriale n. 774 del 4 settembre 2019). <https://www.mim.gov.it/-/linee-guida-dei-percorsi-per-le-competenze-trasversali-e-per-l-orientamento>

Ministero dell'Università e della Ricerca. (n.d.). *PNRR: Misure e componenti*. <https://www.mur.gov.it/it/pnrr/pnrr-misure-e-componenti>

Minutella, V. (2022). Audio description software tools. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 331-352). Routledge.

Molina, L., & Hurtado Albir, A. (2002). Translation techniques revisited: A dynamic and functionalist approach. *Meta*, 47(4), 498-512.

Morisset, L., & Gonant, F. (2008). *Charte de l'audiodescription: Principes et orientations* [online]. <https://www.csa.fr/Media/Files/EspaceJuridique/Chartes/Charte-de-l-audiodescription>

Munday, J. (2016). *Introducing Translation Studies: Theories and applications* (4th ed.). Routledge. <https://doi.org/10.4324/9781315691862>

Murray McGavin, D. D. (1999). The global initiative for the elimination of avoidable blindness - vision 2020: the right to sight. *Community Eye Health*, 12(30), 32.

Nair, J. (2025, January 31). *Man who was born blind explains how he perceives vision and what truly 'blows his mind' about sight*. Ladbible. <https://www.ladbible.com/community/tommy-edison-blind-vision-sense-sight-807955-20250131>

National Endowment for the Arts. (1985). *The Arts and 504: A handbook for accessible arts programming*. U.S. Government Printing Office.

Navarrete, M. (2018). The use of audio description in foreign language education: A preliminary approach. *Translation and Translanguaging in Multilingual Contexts*, 4(1), 129-150. <https://doi.org/10.1075/ttmc.00007.nav>

Navarrete, M. (2020). The use of audio description in foreign language education: A preliminary approach. In L. Incalcaterra McLoughlin, J. Lertola, & N. Talaván (Eds.), *Audiovisual Translation in Applied Linguistics: Educational perspectives* (pp. 131-152). John Benjamins. <https://doi.org/10.1075/bct.111.ttmc.00007.nav>

Navarrete, M. (2022). La audiodescripción como actividad mediadora en el aula de lenguas. In A. Sánchez Cuadrado (Ed.), *Mediación en el aprendizaje de lenguas: Estrategias y recursos* (pp. 41-65). Anaya.

Navarrete, M. (2024). The CEFR CV revisited: Aligning didactic audiovisual translation to the Common European Framework of Reference for Languages. *CEFR Journal - Research and Practice*, 6, 43-55. <https://doi.org/10.37546/JALTSIG.CEFR6-3>

Navarrete, M., & Bolaños García-Escribano, A. (2022). An action-oriented approach to didactic audio description in foreign language education. *Revista de Lenguas para Fines Específicos*, 28(2), 151-164. <https://ojsspdg.ulpgc.es/ojs/index.php/LFE/article/view/1445>

Neves, J. (2005). *Audiovisual translation: Subtitling for the deaf and hard-of-hearing* [Doctoral dissertation] University of Roehampton. https://pure.roehampton.ac.uk/ws/portalfiles/portal/427073/neves_audiovisual.pdf

Nida, E. A., & Taber, C. R. (1982). *The Theory and Practice of Translation*. E. J. Brill. (Original work published 1969)

Nielsen, J., & Loranger, H. (2006). *Prioritizing Web Usability*. New Riders Publishing.

Nord, C. (1997). *Translating as a Purposeful Activity: Functionalist Approaches Explained*. St Jerome.

Nornes, A. M. (2004). For an abusive subtitling. In L. Venuti (Ed.), *The Translation Studies Reader* (pp. 447-469). Routledge. (Reprinted from *Film Quarterly*, 1999, 52(3), 17-34).
<https://doi.org/10.2307/1213822>

Nunan, D. (2006). Task-based language teaching in the Asia context: Defining “task”. *Asian EFL Journal*, 8, 12-18.

Ofcom. (2024). *Online Nation: 2024 Report*.
<https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/online-research/online-nation/2024/online-nation-2024-report.pdf?v=386238>

O'Halloran, K. L., Tan, S., & Marissa, K. L. (2017). Multimodal analysis for critical thinking. *Learning, Media and Technology*, 42(2), 147-170.
<https://doi.org/10.1080/17439884.2016.1101003>

Omvig, J. H. (2009). Why use the word “blind”? *Braille Monitor*, 52(1).
<https://nfb.org/images/nfb/publications/bm/bm09/bm0901/bm090107.htm>

Orero, P. (2005). La inclusión de la accesibilidad en comunicación audiovisual dentro de los estudios de traducción audiovisual. *Quaderns: Revista de traducció*, 10, 173-185.

Orero, P. (2008). Three different receptions of the same film: ‘The Pear Stories Project’ applied to audio description. *Perspectives: Studies in Translatology*, 12(2), 179-193.

Orero, P., & Matamala, A. (2007). Accessible Opera: Overcoming Linguistic and Sensorial Barriers. *Perspectives*, 15(4), 262-277. <https://doi.org/10.1080/13670050802326766>

Orero, P., & Matamala, A. (2013). Standardising audio description. *Italian Journal of Special Education for Inclusion*, 1(1), 150-155.

Orero, P., & Vilaró, A. (2012). Eye tracking analysis of minor details in films for audio description. *MonTI – Monografías de Traducción e Interpretación*, 4, 295-319.
<https://www.redalyc.org/pdf/2651/265125413013.pdf>

Pasqualotto, A., Finucane, C. M., & Newell, F. (2013). Ambient visual information confers a context-specific, long-term benefit on memory for haptic scenes. *Cognition*, 128(3), 363-379.

Pensabene, M. L. (2023). Audio description in English language teaching: Improving language and critical thinking skills. In F. Di Gesù, A. Polizzi, A. Rizzo, & A. Velez (Eds.), *L'accessibilità audiovisiva nelle industrie creative* (pp. 51-73). Palermo University Press.

Perego, E. (2005). *La traduzione audiovisiva*. Carocci Editore.

Perego, E. (2014). *L'audiodescrizione filmica per i ciechi e gli ipovedenti*. Trieste: EUT.

Perego, E. (2016). Gains and losses of audio description in sighted viewers. *Target*, 28(3), 424-444.

Perego, E. (2017). Norms for audio description in Italy: The case of Senza Barriere. *RITT (Rivista Internazionale di Tecnica della Traduzione)*, 19, 207-228.

Perego, E. (2018). Cross-national research in audiovisual translation: Some methodological considerations. *Journal of Audiovisual Translation*, 1(1), 63-80.

Perego, E. (2020). *Accessible Communication: A Cross-country Journey*. Frank & Timme.
https://iris.unipv.it/retrieve/faeb62c6-7ecb-4d37-85f1-d480c4ed2826/A_90654_Layout_3.pdf

Perego, E., & Taylor, C. (2012). *Tradurre l'audiovisivo*. Carocci Editore.

Pérez-Ramos, A. (1988). *Francis Bacon's Idea of Science and the Maker's Knowledge Tradition*. Oxford University Press.

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/10.1177/0145482X04098008>

Prado Martinez, R. (2015, August 24). *Film Vs. Musical: Billy Elliot*. <https://ricardopradamartinez.wordpress.com/2015/08/24/film-vs-musical-billy-elliott/>

Psyché, V., Daniel, B. K., & Bourdeau, J. (2020). Learning spaces in context-aware educational networking technologies in the digital age. In A. Peña-Ayala (Ed.), *Educational Networking. Lecture Notes in Social Networks*. Springer, Cham. https://doi.org/10.1007/978-3-030-29973-6_9

Pujiyanti, I., & Ikaputra, I. (2025). Universal design is “not only” a design for people with disabilities. *Journal of Architectural Research and Design Studies*, 9(1), 48-55.

Pujol, J., & Orero, P. (2007). Audio description precursors: Ekphrasis and narrators, *Translation Watch Quarterly*, 3(2), 49-60.

Pym, A. (2003). Redefining Translation Competence in an Electronic Age. In Defence of a Minimalist Approach. *Meta*, 48(4), 481-497. <https://doi.org/10.7202/008533ar>

Rai, S., Greening, J., & Petré, L. (2010). *A comparative study of audio description guidelines prevalent in different countries*. RNIB.

Ramos, M. (2015). The emotional experience of films: does audio description make a difference? *The Translator*, 21(1), 68-94. <https://doi.org/10.1080/13556509.2014.994853>

Ranzato, I. (2014). Period television drama: culture specific and time specific references in translation for dubbing. In M. Pavesi, M. Formentelli, & E. Ghia (Eds.), *The Languages of Dubbing: Mainstream Audiovisual Translation in Italy*. Peter Lang Verlag.

Reiss, K. (2000). *Translation Criticism: Potential and Limitations* (E. F. Rhodes, Trans.). St Jerome and American Bible Society. (Original work published 1971)

Remael, A., & Vercauteren, G. (2007). Audio describing the exposition phase of films: Teaching students what to choose. *Trans: Revista de Traductología*, 11, 73-93.

Remael, A., & Vercauteren, G. (2010). The translation of recorded audio description from English into Dutch. *Perspectives*, 18(3), 155-171.

Remael, A., & Vercauteren, G. (2011). *Basisprincipes voor audiobeschrijving voor televisie en film* [Basics of audio description for television and film]. Departement Vertalers and Tolken, Artesis Hogeschool.

Remael, A., & Vercauteren, G. (2015). Spatio-temporal settings and their continuity. In A. Remael, N. Reviers, & G. Vercauteren (Eds.), *Pictures painted into words: ADLAB Audio description guidelines* (pp. 24-27). Trieste:EUT.

Remael, A., & Reviers, N. (2019). Multimodality and audiovisual translation: Cohesion in accessible films. In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation* (pp. 260-280). Routledge.

Remael, A., Orero, P., Black, S., & Jankowska, A. (2019). From translators to accessibility managers: How did we get there and how do we train them? In M. Tolosa Igualada & Á. Echeverri (Eds.), *Porque algo tiene que cambiar. La formación de traductores e intérpretes: Presente & futuro / Because something should change: Present & Future Training of Translators and Interpreters* (pp. 131-154). MonTI 11.

Rizzo, A. (2025). Creativity in audio description: Culture-specific references in I leoni di Sicilia. *Translation Spaces*, 14(2). <https://doi.org/10.1075/ts.25051.riz>

Rizzo, A., & Pensabene, M. L. (2021). Accessibilità e contro narrazione: Il Sole Luna Doc Festival. *Lingue e Linguaggi*, 43, 85-118. <https://doi.org/10.1285/i22390359v43p85>

Rizzo, A., & Spinzi, C. G. (2023). Authorial (audio) description: Creativity in the transfer of CSRs in Squid Game. *Translation and Translanguaging in Multilingual Contexts*, 9(3), 419-449. <https://doi.org/10.1075/ttmc.00121.riz>

Rodrigues Barbosa, E. (2013). La audiodescripción como herramienta para la enseñanza de ELE. In L. M. A. de Freitas (Ed.), *Anais do XIV Congresso Brasileiro de Professores de Espanhol: Língua e Ensino 1* (pp. 487-496). APEERJ. <http://apeerj.blogspot.it/2013/07/anais-do-xiv-congresso-brasileiro-de.html>

Romero, L. (2015). L'uso del doppiaggio e del sottotitolaggio nell'insegnamento della L2: Il caso della piattaforma ClipFlair. *Lingue e Linguaggi*, 15, 223-242. DOI 10.1285/i22390359v15p277

Romero-Fresco, P. (2013). Joining the dots between audiovisual translation, accessibility and filmmaking. *The Journal of Specialised Translation*, 20, 201-223. <https://doi.org/10.26034/cm.jostrans.2013.410>

Romero-Fresco, P. (2017a). Accessible filmmaking in documentaries. *inTRAlinea*, Special Issue: Building Bridges between Film Studies and Translation Studies. <https://www.intraline.org/specials/article/2251>

Romero-Fresco, P. (2017b, September 25-26). *Widening the scope of media accessibility* [Conference presentation]. Intermedia: International Conference on Audiovisual Translation, Poznań, Poland.

Romero-Fresco, P. (2018). In support of a wide notion of media accessibility: Access to content and creation. *Journal of Audiovisual Translation*, 1(1), 187-204. <https://doi.org/10.47476/jat.v1i1.53>

Romero-Fresco, P., & Fryer, L. (2019). *Accessible filmmaking guide*. BFI.

Romero-Fresco, P. (2019a). Accessible filmmaking: Translation and accessibility from production. In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation* (pp. 498-515). Routledge.

Romero-Fresco, P. (2019b). *Accessible Filmmaking: Integrating translation and accessibility into the filmmaking process*. Routledge.

Romero-Fresco, P. (2019c). Training in accessible film-making. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 47-72.
<https://doi.org/10.52034/lanstts.v18i0.513>

Romero-Fresco, P., & Chaume, F. (2022). Creativity in audiovisual translation and media accessibility. *The Journal of Specialised Translation*, 38, 75-101.
<https://doi.org/10.26034/cm.jostrans.2022.084>

Romero-Fresco, P., & Dangerfield, K. (2022). Accessibility as a conversation. *Journal of Audiovisual Translation*, 5(2), 15-34. <https://doi.org/10.47476/jat.v5i2.2022.228>

Romero-Muñoz, A. (2023). Multimodal analysis as a way to operationalise objectivity in audio description: A corpus-based study of Spanish series on Netflix. *Journal of Audiovisual Translation*, 6(2), 8-32. <https://doi.org/10.47476/jat.v6i2.2023.251>

Romero-Muñoz, A. (2025). Exploring the accessibility creative continuum on streaming platforms: A contrastive multimodal analysis of subjectivity and objectivity in audio description. *The International Journal of Translation & Interpreting Research*, 17(2), 155-172.

Rothery, J., & Stenglin, M. (2000). Interpreting literature: The role of APPRAISAL. In L. Unsworth (Ed.), *Researching Language in Schools and Functional Linguistic Perspectives*. Cassell.

Roulstone, A. (2016). *Disability and Technology: An Interdisciplinary and International Approach*. Palgrave Macmillan.

Royal National Institute of Blind People. (2024, November 6). *Blind content creator and disability advocate Claire Sisk announced as RNIB Ambassador* [Press release]. <https://www.rnib.org.uk/news/blind-content-creator-and-disability-advocate-claire-sisk-announced-as-rnib-ambassador/>

Sagar, S. (2000, September 19). Billy Elliot story 'still growing' 25 years on. *BBC News*. <https://www.bbc.com/news/articles/cx2xvd84xqvo>

Sanz-Moreno, R. (2017). La (auto)censura en audiodescripción. El sexo silenciado. *Parallèles*, 29(2), 46-63.

Sanz-Moreno, R. (2017). The audio describer as a cultural mediator. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, 30(2), 538-558. <https://doi.org/10.1075/resla.00006.san>

Sanz-Moreno, R. (2018). El uso de extranjerismos en audiodescripción. La opinión de los usuarios. *Parallèles*, 30(2), 53-69.

Schaeffer-Lacroix, E., & Schulz, S. J. (2021, September 30-October 1). *Creating and translating audio descriptions with the help of Frazier* [Conference presentation]. Interdisciplinary Conference "Vers une robotique du traduire? [Towards a Robotics of Translation?]", Université de Strasbourg. <https://hal.science/hal-03905096v1>

Schjoldager, A. G., Gottlieb, H., & Klitgård, I. (2008). *Understanding Translation*. Systime Academic.

Schleppegrell, M. J., & Oteíza, T. (2023). Systemic functional linguistics: Exploring meaning in language. In J. P. Gee & M. Handford (Eds.), *The Routledge Handbook of Discourse Analysis* (2nd ed., pp. 156-169). Routledge. <https://doi.org/10.4324/9781003035244-13>

Shakespeare, T. (2010). The social model of disability. In J. D. Lennard (Ed.), *The Disability Studies Reader* (266-273). Routledge.

Snyder, J. (2007). Audio description: The visual made verbal. *The International Journal of the Arts in Society: Annual Review*, 2(2), 99-104. <https://doi.org/10.18848/1833-1866/CGP/v02i02/35358>

Snyder, J. (2014). *The Visual Made Verbal: A Comprehensive Training Manual and Guide to the History and Applications of Audio Description*. American Council of the Blind.

Snyder, J. (2016, August 11-12). *Literacy and audio description* [Conference presentation]. IFLA Libraries for Print Disabilities (LPD) Satellite Conference: Tailoring the Reading Experience to Meet Individual Needs, Louisville, KY. <https://www.ifla.org/wp-content/uploads/2019/05/assets/libraries-for-print-disabilities/conferences-seminars/ad-ifla-2016-literacy-and-audio-description.pdf>

Snyder, J. (2022). Audio description in the United States. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 531-543). Routledge.

Snyder, J. (2026, April 23-24). *AD (Audio Description) and AI: For good or ill?* [Conference presentation]. Artificial Intelligence and Audiovisual Translation, University of Palermo.

Snyder, J. (n.d.). *Audio Description: An International Phenomenon*. American Council of the Blind. Retrieved November 23, 2025, from <https://www.acb.org/audio-description-international-phenomenon>

Sokoli, S. (2006). Subtitling for the deaf and hard of hearing: The impact of new technologies. *Proceedings of the 2006 International Conference on Media Literacy*. Retrieved July 17, 2026, from https://www.euroconferences.info/proceedings/2006_Proceedings/2006_Sokoli_Stravoul_a.pdf

Soler Gallego, S., & Luque Colmenero, M. O. (2023). Increased subjectivity in audio description of visual art: a focus group reception study of content minimalism and interpretive voicing. *Journal of Audiovisual Translation*, 6(2), 55-76.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed.). Blackwell.

Spinzi, C. (2016). *Discursive Strategies in the Language of Foreign Policy: The Semantics and the Appraisal of Security*. Wip Edizioni.

Spottiswoode, R. (1950). *A Grammar of the Film: An Analysis of Film Technique*. University of California Press.

Stephanidis, C., & Emiliani, P. L. (1999). Connecting to the Information Society: a European perspective, *Technology and Disability*, 10(1), 21-44. <https://doi.org/10.3233/TAD-1999-10>

Sterrit, D. (2000, October 13). Working-class British tales leap onto screens. *The Christian Science Monitor*. <https://www.csmonitor.com/2000/1013/p15s2.html>

Switzer, B. (2021, June 16). *Evolution and inclusive language: Creating a new term for visual disabilities*. Blind New World. <https://blindnewworld.org/evolution-and-inclusive-language-creating-a-new-term-for-visual-disabilities/>

Szarkowska, A. (2011). Text-to-speech audio description: Towards wider availability of AD, *The Journal of Specialised Translation*, 15, 142-162. <https://doi.org/10.26034/cm.jostrans.2011.509>

Szarkowska, A. (2013). Auteur description: From the director's creative vision to audio description. *Journal of Visual Impairment & Blindness*, 107(5), 383-387.

Szarkowska, A., & Jankowska, A. (2015). Audio describing foreign films. *The Journal of Specialised Translation*, 23, 243-269.

Szarkowska, A., Krejtz, I., Krejtz, K., & Duchowski, A. (2013). Harnessing the potential of eyetracking for media accessibility. In S. Grucza, M. Płużyczka, & J. Zając (Eds.), *Translation Studies and Eye-tracking Analysis* (pp. 153-183). Peter Lang.

Tait, K., & Silveira, S. (2023). Easing the tension between medical and social models of disability: the biopsychosocial model of disability. *InSpEd Insights*. <https://www.insped.org.au/wp-content/uploads/2023/12/Easing-the-Tension-between-Medical-and-Social-Models-of-Disability-The-Biopsychosocial-Model-of-Disability-Dec-23.pdf>

Talaván, N. (2010). Subtitling as a task and subtitles as support: Pedagogical applications. In J. Díaz Cintas, A. Matamala, & J. Neves (Eds.), *New insights into audiovisual translation and media accessibility* (pp. 285-299). Rodopi.

Talaván, N. (2020). The didactic value of AVT in foreign language education. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 567-592). Palgrave Macmillan.

Talaván, N., & Lertola, J. (2016). Active audiodescription to promote speaking skills in online environments. *Sintagma*, 28, 59-74. <https://doi.org/10.21001/sintagma.2016.28.04>

Talaván, N., & Lertola, J. (2022). Audiovisual translation as a didactic resource in foreign language education: A methodological proposal. *Encuentro*, 30, 23-39.

Talaván, N., & Lertola, J. (2025). Creativity in didactic audiovisual translation: Possibilities and preliminary findings. *Translation Spaces*. <https://doi.org/10.1075/ts.25016.ta>

Talaván, N., & Rodríguez-Arancón, P. (2014). The use of reverse subtitling as an online collaborative language learning tool. *The Interpreter and Translator Trainer*, 8(1), 84-101. <https://doi.org/10.1080/1750399X.2014.908559>

Talaván, N., Lertola, J., & Costal, T. (2016). iCap: Intralingual captioning for writing and vocabulary enhancement. *Alicante Journal of English Studies*, 29, 229-248. <https://doi.org/10.14198/raei.2016.29.13>

Talaván, N., Lertola, J., & Fernández-Costales, A. (2024). *Didactic Audiovisual Translation and Foreign Language Education*. Taylor & Francis.

Taylor, C. (2003). Multimodal transcription in the analysis, translation and subtitling of Italian films. *The Translator*, 9(2), 191-205. <https://doi.org/10.1080/13556509.2003.10799153>

Taylor, C. (2015). The language of AD. In A. Remael, N. Reviers, & G. Vercauteren (Eds.), *Pictures painted into words: ADLAB Audio Description guidelines* (pp. 50-54). Trieste: EUT.

Taylor, C. (2016). The multimodal approach in audiovisual translation. *Target: International Journal of Translation Studies*, 28(2), 222-236. <https://doi.org/10.1075/target.28.2.04tay>

Taylor, C. (2020). Multimodality and intersemiotic translation. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (pp. 83-100). Palgrave Macmillan.

Taylor, C., & Perego, E. (Eds.). (2022). *The Routledge Handbook of Audio Description*. Routledge.

Taylor Torsello, C. (1996). On the logical metafunction. *Functions of Language*, 3(2), 151-183. <https://doi.org/10.1075/fol.3.2.02tor>

Thompson, G. (2014). *Introducing Functional Grammar*. Routledge. <https://doi.org/10.4324/9780203785270>

Thompson, H. (2015, August 3). *How the arts can help change attitudes to blindness*. The Guardian. <https://www.theguardian.com/culture-professionals-network/2015/aug/03/arts-change-attitude-blindness-blind-creations>

Tianli, Z., Mansor, N. S., Ang, L. H., & Sharmini, S. (2022). Does Ellen DeGeneres adopt rhetorical strategies in her talk show's monologues for verbal humor? *Journal of Language and Linguistic Studies*, 18(1), 249-265.

Titchkosky, T. (2011). *The Question of Access: Disability, Space, Meaning*. University of Toronto Press.

Toffler, A. (1980). *The Third Wave*, New York, William Morrow and Company, INC.

Truby, J. (2007). *The Anatomy of Story: 22 steps to becoming a master storyteller*. Faber & Faber.

Tuominen, T., Koponen, M., Vitikainen, K., Sulubacak, U., & Tiedemann, J. (2023). Exploring the gaps in linguistic accessibility of media: The potential of automated subtitling as a solution. *The Journal of Specialised Translation*, 39, 77-98. <https://doi.org/10.26034/cm.jostrans.2023.067>

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>

Udo, J. P., & Fels, D. I. (2010). Universal design on stage: Live audio description for theatrical performances. *Perspectives: Studies in Translatology*, 18(3), 189-203. <https://doi.org/10.1080/0907676X.2010.485683>

Udo, J. P., Fels, D. I., Ting, P., Diamond, J. E., & Diamond, J. I. (2005). Odd Job Jack Described – A first person narrative approach to described video. *Journal of Universal Access in the Information Society*, 5(1), 73-81.

Unione Italiana dei Ciechi e degli Ipovedenti (UICI) Sezione di Verona. (n.d.). *Supporto tecnologico. Servizi e Attività*. Retrieved July 17, 2026, from <https://www.uiciverona.it/servizi-e-attivita/supporto-tecnologico/>

Uzzo, G. (2024). *Accessibility at film festivals: Guidelines for inclusive subtitling* [Unpublished doctoral dissertation] University of Palermo.

Vercauteren, G. (2007). Towards a European guideline for audio description. In J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for all: Subtitling for the deaf, audio description, and sign language* (pp. 139-149). Rodopi.

Vercauteren, G. (2012). A narratological approach to content selection in audio description: Towards a strategy for the description of narratological time. *MonTI: Monografías de Traducción e Interpretación*, 4, 207-231. <https://doi.org/10.6035/MONTI.2012.4.9>

Vercauteren, G. (2016). *A narratological approach to content selection in audio description* [Unpublished doctoral dissertation] University of Antwerp.

Vercauteren, G. (2017). Audio describing the extradiegetic emotive dimension of film. *International Journal of Translation*, 19, 163-182.

Vercauteren, G. (2021). Insights from mental model theory and cognitive narratology as a tool for content selection in audio description. *Journal of Audiovisual Translation*, 4(3), 6-24. <https://doi.org/10.47476/jat.v4i3.2021.191>

Vercauteren, G. (2022). Narratology and/in audio description. In C. Taylor & E. Perego (Eds.), *The Routledge Handbook of Audio Description* (pp. 78-92). London and New York: Routledge.

Vercauteren, G., & Orero, P. (2013). Describing facial expressions: much more than meets the eye. *Quaderns: Revista de Traducció*, 20, 187-199. <https://ddd.uab.cat/record/107207>

Vinay, J. P., & Darbelnet, J. (1995). *Comparative Stylistics of French and English: A methodology for translation*. John Benjamins.

Vincendeau, G. (1999). Hollywood Babel: The coming of sound and the multiple-language version. In A. Higson & R. Maltby (Eds.), *"Film Europe" and "Film America": Cinema, Commerce and Cultural Exchange 1920-1939* (pp. 207-225). University of Exeter Press.

Vygotsky, L. S., & Cole, M. (1978). *Mind in Society: Development of higher psychological processes*. Harvard University Press.

Walczak, A. (2017). Creative description: Audio describing artistic films for individuals with visual impairments. *Journal of Visual Impairment & Blindness*, 111(4), 387-391.

Walczak, A., & Fryer, L. (2017a). Creative description: The impact of audio description style on presence in visually impaired audiences. *British Journal of Visual Impairment*, 35(1), 6-17. <https://doi.org/10.1177/02646196166616>

Walczak, A., & Fryer, L. (2017b). Vocal delivery of audio description by genre: measuring users' presence. *Perspectives*, 26(1), 69-83. <https://doi.org/10.1080/0907676X.2017.1298634>

Walczak, A., & Szarkowska, A. (2012). Text-to-speech audio description to educational materials for visually-impaired children. In S. Bruti & E. Di Giovanni (Eds.), *Audiovisual Translation across Europe: An ever-changing landscape* (pp. 209-234). Peter Lang.

Wang, B., & Ma, Y. (2021). *Systemic Functional Translation Studies: Theoretical Insights and New Directions*. Equinox.

Wang, B., & Ma, Y. (2023). *Theorizing and Applying Systemic Functional Linguistics: Developments by Christian M.I.M. Matthiessen*. Routledge. <https://doi.org/10.4324/9781003041238>

White, P. R. R. (2022). The Appraisal framework and analyses of journalistic discourse: Objectivity, subjectivity and attitudinal positioning. In D. Caldwell, J. S. Knox, & J. R. Martin (Eds.), *Applicable Linguistics and Social Semiotics: Developing Theory From Practice* (pp. 295-308). Bloomsbury Academic.

Wilken, N., & Kruger, J. (2016). Putting the audience in the picture: Mise-en-shot and psychological immersion in audio described film. *Across Languages and Cultures*, 17(2), 251-270.

World Health Organization. (1980). *International Classification of Impairments, Disabilities, and Handicaps: A manual of classification relating to the consequences of disease*. <https://iris.who.int/server/api/core/bitstreams/50556222-5563-44a8-b0db-fd08c6943f29/content>

World Health Organization. (2001). *International Classification of Functioning, Disability and Health*. <https://iris.who.int/server/api/core/bitstreams/83e1c3e0-cf9e-4063-a424-71ccc9a84ded/content>

World Health Organization. (2022). *International Classification of Diseases 11th Revision*. <https://icd.who.int/en/>

World Health Organization. (2026, February 10). *Blindness and vision impairment*. <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>

World Wide Web Consortium. (2023). *Web content accessibility guidelines (WCAG) 2.2*. <https://www.w3.org/TR/WCAG22/>

Zabrocka, M. (2019). Speaking metaphor in audio description for children. *Culture, Language and Representation*, 22, 153-169.

Zanotti, S. (2022). Audiovisual translation. In K. Malmkjaer (Ed.), *The Cambridge handbook of translation* (pp. 440-460). Cambridge University Press.

Zanotti, S., & Sánchez-Mompeán, S. (2025). Introduction. Riding the streaming wave: Broadening global access through audiovisual translation and media accessibility. *The International Journal of Translation and Interpreting Research*, 17(2), 1-15.

Zhang, E. D., & Yu, S. (2024). *Assessing L2 Digital Multimodal Composing Competence* (1st ed.). Routledge. <https://doi.org/10.4324/9781003475729>

Sitography

<https://www.aao.org/>
<https://www.acb.org/home>
<https://www.adlabproject.eu>
<https://www.aenor.com>
<https://www.alteff2020.eventive.org/films/5f9d82115c955f02c39bfd07>
<https://www.artis-project.com/>
<https://www.bbc.com/>
<https://www.bfi.org.uk/>
<https://www.clipflair.net>
<https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions>
<https://www.dubaifilmfest.com/>
<https://www.euroblind.org>
www.gradesaver.com
<https://www.grammatics.com/appraisal/>
<https://www.iapb.it>
<https://www.iapb.org>
<https://www.icevi.org>
<https://www.laurenceanthony.net>
<https://www.literarylattitude.com>
<https://www.makethingsaccessible.com>
<https://www.mediaforall.eu/>
<https://www.mim.gov.it/>
<https://www.mur.gov.it>
<https://www.nfb.org>
<https://www.nipai.org>
<https://www.olympics.com/it/olympic-games/milano-cortina-2026>
<https://www.psicologo-milano.it/newblog/psicologia-cecita-vista/>
<https://www.rai.it>
<https://www.rnib.org.uk/>
<https://www.senzabarriere.org/>
<https://www.solelunadoc.org>
<https://www.tommyedison.com>
<https://www.tradilex.es/en>
<https://webs.uab.cat/arsad/>
<https://www.uiciechi.it/>
<https://www.videotovoice.com>
<https://www.w3.org>
<https://www.wbu.ngo>
<https://www.who.int>
<https://www.youtube.com>

Filmography

- Besson, L. (Director). (1999). *The Messenger. The Story of Joan of Arc* [Film].
- Chanan, M. (Director). (2012). *Secret City* [Film].
- Cole, A. (Director). (2016). *The Colours of the Alphabet* [Film].
- Coppola, F.F. (1988). *Tucker: The Man and His Dream* [Film].
- Crossland, A. (Director). (1927). *The Jazz Singer* [Film].
- Daldry, S. (Director). (2000). *Billy Elliot* [Film].
- Holbole, R. L. (Director). (2018). *Aagaswadi -Village in the Sky* [Short Film].
- Middleton, P. & Spinney, J. (Directors). (2016). *Notes on Blindness* [Film].
- Patrick, L. (Director). (2003-2022). *The Ellen DeGeneres Show* [Talk Show].
- Pêra, E. (Director). (2011). *O Barão (The Baron)* [Film].
- Romero-Fresco, P. (Director). (2012). *Joining the Dots* [Short Film].